

ABSTRACT

Title of Dissertation: FUNCTIONAL STUDY OF SAPOSIN-LIKE
PROTEINS IN *ARABIDOPSIS THALIANA*

Changxu Pang, Doctor of Philosophy, 2020

Dissertation Directed By: Dr. Angus Murphy, Professor, Department of
Plant Science & Landscape Architecture

Dr. Wendy Ann Peer, Associate Professor,
Department of Environmental Science &
Technology

Sphingolipids and microdomain-associated proteins that are associated with the plasma membrane and endomembrane system are important in plant growth and development. Elucidating functions of these proteins advance understanding of signal transduction from plasma membrane into cytosol and between different intracellular membrane compartments. Saposins and saposin-like proteins (SAPLIP) are among these proteins. In plants, two types of proteins contain saposin B-like domains (SapB-like domains): saposin-domain containing aspartic proteases (ASPAs) and prosaposin-like proteins (PSAPLIPs).

Phenotypic analyses showed that single loss-of-function *aspa2* showed delayed seed maturation. Seeds of *aspa1-2 aspa2-1 aspa3-3* triple mutant (*aspa1* is knock-

down, *aspa2* and *aspa3* are knock-out alleles) showed delayed germination rates and delayed seed storage proteins degradation. Further, protein storage vacuolar fusion was also delayed in the mutant cotyledons. These results suggest that ASPAs process seed storage proteins during seed germination *in vivo*, and probably also involved in protein storage vacuolar fusion regulation.

ASPAs also have a role in root architecture. Triple mutant showed longer primary root length under low nitrogen conditions. Further analysis suggested that the altered root architecture in the mutants may result from rates of tracheary element (TE) maturation in xylem tissues. Triple mutants were slightly delayed in TE maturation and the *ASPA2* overexpression showed slightly early maturation. Together with the expression pattern of *ASPA3*, this indicates that ASPAs may take part in programmed cell death (PCD) in *Arabidopsis*. Further studies showed that ASPAs are involved in PCD execution. Results showed that the onset of PCD was not delayed in the triple mutant, but the execution time of PCD was extended. Membrane permeability increased more slowly in the triple mutants and faster in the overexpression plants. This reflects the role of ASPAs in membrane disturbance and permeability regulation during PCD.

The prosaposin-like proteins (PSAPLIPs) have received little study. Sequence alignments identified that prosaposin-like proteins are ubiquitous in plant kingdom. Plant PSAPLIPs show highly conserved in secondary structure of SapB-like domains. This structural similarity was supported by glycosylation analyses of *Arabidopsis thaliana* *AtPSAPLIP1* and *AtPSAPLIP2*. Both *AtPSAPLIP1* and *AtPSAPLIP2* traffic to

vacuoles. Possible role of PSAPLIPs is facilitating target protein degradation. *AtPSAPLIP1* was mainly expressed in inflorescence, especially in sepals, carpels and mature pollen grains, as well as leaves and roots. Young leaves had higher expression level than aged leaves. *AtPSAPLIP2* was expressed in inflorescence too, but mainly in young anthers, petals, ovules and developing seeds. This result indicates function differentiation of PSAPLIPs in *Arabidopsis*. Both genes are important in male gametophyte development.

The significance of this dissertation is that it demonstrates that ASPAs process seed storage proteins during seed germination *in vivo* for the first time. It also discovered a new role of ASPAs in regulating programmed cell death by promoting memberane permeability, and thus affecting root growth in *Arabidopsis*. The third is that this is also the first time to characterize the plant prosaposin-like proteins, which are important in male gametophyte development and provide novel sights on how plants regulate reproductive process. These results will broaden our understanding of the protein-lipid interaction in the cell and the biological functions of saposin-like proteins in plant growth and development.

FUNCTIONAL STUDY OF SAPOSIN-LIKE PROTEIN IN *ARABIDOPSIS*
THALIANA

by

Changxu Pang

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2020

Advisory Committee:

Professor Angus Murphy, Chair

Prof. Wendy Peer, Co-Chair

Prof. Caren Chang

Prof. Gary Coleman

Prof. Zhongchi Liu

© Copyright by
Changxu Pang
2020

Preface

This dissertation is original, unpublished, independent work by the author, Changxu Pang. I am responsible for all major areas of data collection and analysis, and majority of figures. Ally Albers and Amani Perwaz Aulakh helped in genotyping the plants.

This dissertation is composed of four chapters and five appendices. Each chapter is organized in a manuscript format (abstract, introduction, results, discussion, materials and methods). Materials and methods listed in the end of each chapter are used in this chapter only, and details in materials and methods are listed in appendix D and some are repeated. Supplementary figures and tables are included in the appendices. Supplementary figures for chapter 2 are in appendix A. Supplementary figures for chapter 3 are in appendix B and appendix C. Primer list and plant prosaposin-like protein gene list are in appendix E. Additional figures in other projects during PhD are listed in appendix F.

I would like to give acknowledgement to Dr. Liwen Jiang who provided the GFP-FREE1 marker line in this dissertation.

Acknowledgements

First, I would like to show my gratitude to my advisor, Dr. Angus Murphy, for giving me the opportunity to work in this lab for all these years. He has been quite supportive both in this project and in my personal life. He is patient with all the questions that I asked and provides informative ideas for me. Every time that I encounter a hard situation, he encouraged me to overcome difficulties. I'm glad that I received academic trainings in this lab. Besides, I wish to pay my regards to my co-advisor Dr. Wendy Peer for academic advice (and cakes). Conversations with her have always been glad and helpful.

Then I wish to express my sincere gratitude to the rest of my committee members, Dr. Caren Chang, Dr. Zhongchi Liu and Dr. Gary Coleman, for your perspectives on my project and great advice.

I would also give my thanks to my colleagues Dr. Mark Jenness, Dr. Candace Prichard, Dr. Jun Zhang, Dr. Rueben Tayengwa, Dr. Doron Shkolnik, Dr. Wiebke Tapken, Sarah Turner, Ally Albers, Amani Perwaz Aulakh, Gabreille Bate and Juliane Henschel. I'm really appreciated for your advice whenever I needed and keep the lab operations running very well. Without your help, I could not go through all the work. I would like to recognize the invaluable assistance which you provided during my study.

And last, I would like to express my greatest gratitude to my mother. Though we are apart for a long time, I still feel her deepest love and care, which is my greatest

motivation for this venture.

Table of Contents

Preface	ii
Acknowledgements.....	iii
Abbreviations.....	viii
List of Figures	x
Chapter 1: Literature review: Saposin-like proteins in plants	1
Abstract	1
Introduction: Sphingolipids and microdomains in plants	1
Saposin-like Proteins (SAPLIPs)	5
Biological functions of SAPLIPs in animals.....	6
SAPLIPs in plants.....	8
General structure features of SAPLIPs	10
Structure and function of human saposins.	11
Structure of and function of prosaposins.....	11
Structure and function of mature saposins.....	13
Structural features of saposin A.....	15
Structure feature of saposin C.....	15
Structure feature of saposin B.....	16
Structure feature of saposin D	17
Mechanistic model of saposin-lipid interactions	18
Structure of SAPLIPs in plants	21
Plant aspartic proteases	23
General Information about plant aspartic protease	23
Aspartic proteases in <i>Arabidopsis</i>	24
Biological functions of aspartic proteases in plants	27
Structure of plant aspartic protease	33
Biological function of plant specific insert	35
Perspective	38
Chapter 2 Functional study of saposin-like domain containing aspartic proteases in <i>Arabidopsis thaliana</i>	40
Abstract	40
Introduction.....	41

Results	47
ASPAs function in seed development and germination	47
Expression pattern of <i>ASPA2</i>	58
Subcellular localization and trafficking of <i>ASPA2</i>	62
ASPAs are involved in root architecture regulation	67
Transcriptional regulation of <i>ASPA2</i>	70
ASPAs are involved in programmed cell death	74
Discussion	80
Conclusion	87
Materials and Methods	89
Plant materials	89
Germination test	90
Seed Protein extraction	91
SDS-PAGE	91
Coomassie blue staining	92
Glycosylation test	92
Western blot	92
Cloning and expression vector construction	93
Microscopy	94
Time-course image of PI (propidium iodide) and PI/FDA (fluorescein diacetate) double staining in lateral root cap	95
Chapter 3: Elucidating features and functions of plant prosaposin-like proteins (PSAPLIPs)	96
Abstract	96
Introduction	97
Results	101
Phylogenic studies of PSAPLIPs in plants	101
Structural features of AtPSAPLIPs	106
Subcellular localization of <i>Arabidopsis</i> PSAPLIPs	124
Expression pattern of <i>AtPSAPLIP1</i> and <i>AtPSAPLIP2</i>	128
Discussion	132

Conclusion	137
Materials and Methods	138
Primary and Secondary Structure Prediction	138
Sequence Alignment.....	139
Phylogenetic tree construction	140
Preparation of Transgenic Plants	140
Plant Materials and chemical treatment.....	141
Protein extraction.....	142
Glycosylation test	142
SDS-PAGE.....	143
Western blot.....	143
Microscopy	144
Histochemistry.....	145
Chapter 4: Conclusion and Perspective	146
Appendix A Supplemental Figures for Chapter 2	153
Appendix B Supplemental Figures for Chapter 3.....	160
Appendix C Phylogenic tree of plant PSAPLIPs	216
Appendix D Materials and Methods	224
Appendix E Supplemental Tables.....	242
Appendix F Additional Data	274
References	292

Abbreviations

Abbreviation and symbol	Definition
A	alanine
AA	amino acid
ABA	abscisic acid
Amp	ampicillin
ASPA	aspartic protease
<i>aspa1-2/2-1/3-3</i>	<i>aspa1-2 aspa2-1 aspa3-3</i> triple mutant
BFA	brefeldin A
bp	base pairs
Col-0	Columbia-0
CFP	cyan fluorescent protein
Conc A	concanamycin A
CRISPR	clustered regularly interspaced short palindromic repeats
CTAB	cetyl trimethylammonium bromide
D	aspartic acid
DAG	days after germination
DAP	days after pollination
DMSO	Dimethyl sulfoxide
DPI	diphenyleneiodonium
DNA	deoxyribonucleic acid
EDTA	ethylenediaminetetraacetic acid
E	glutamic acid
<i>E. coli</i>	Escherichia coli
EE	early endosome
FDA	fluorescein diacetate
GA	gibberellic acid
Gent	gentamicin
GFP	green fluorescent protein
GUS	β-glucuronidase
H ₂ A	histone 2A 10
H ₂ O ₂	hydrogen peroxide
HCl	hydrochloric acid
Hyg	hygromycin
K	lysine
Kan	Kanamycin
kb	kilo base pairs
kDa	kilo Dalton
KNO ₃	potassium nitrate
L	leucine
LB	Luria-Bertani

LE	late endosome
MS	Murashige and Skoog
MVB	multivesicular body
N	asparagine
NaCl	sodium chloride
NADPH	nicotinamide adenine dinucleotide phosphate
NaOH	sodium hydroxide
NTPP	N-terminal propeptide
P	proline
PBS	phosphate-buffered saline
PCD	programmed cell death
PCR	polymerase chain reaction
PI	propidium iodide
PM	plasma membrane
PSAPLIP	prosaposin-like protein
PSI	plant specific insert
PSV	protein storage vacuole
PVC	prevacuolar compartment
PVDF	Polyvinylidene fluoride
Q	glutamine
R	arginine
Rif	rifampicin
RFP	red fluorescent protein
RNA	ribonucleic acid
rpm	round per minute
SapB	saposin B
SDS	sodium dodecyl sulfate
SDS PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SAPLIP	saposin-like protein
SP	signal peptide
Spec	spectinomycin
SSP	seed storage protein
TE	tracheary element
TEMED	Tetramethyl ethylenediamine
TGN	<i>trans</i> -Golgi network
TML	transmitted light
WT	wild type
X gluc	5-Bromo-4-chloro-3-indolyl- β -D-glucuronide
YFP	yellow fluorescent protein
Zeo	zeocin

List of Figures

Figure 1-01. Schematic depiction of the three main activities of SAPLIPs.

Figure 1-02. Schematic comparison between animal aspartic proteases and plant aspartic proteases.

Figure 1-03. Schematic comparison between saposin and swaposin.

Figure 1-04. Schematic illustration of saposin domain of NK-lysin and the swaposin domain of prophytepsin from barley.

Figure 1-05. Models of lipid activation by saposins.

Figure 1-06. A model for the interaction with membranes and a general mechanism of action of the SAPLIP family.

Figure 1-07. Plant PSI primary structure.

Figure 1-08. Phylogenetic tree of A1 family proteases in *Arabidopsis*.

Figure 1-09. Schematic illustration of the process steps for ASPA1.

Figure 2-01. Characterization of *ASPA* T-DNA insertion mutants.

Figure 2-02. *ASPAs* regulate seed maturation in *Arabidopsis*.

Figure 2-03. Seed germination and seed storage protein degradation in *aspa* mutant.

Figure 2-04. Germination rates in Col-0 and *aspa* triple mutant seeds with and without gibberellin acid 3, hydrogen peroxide and diphenylene iodonium treatments.

Figure 2-05. Protein storage vacuole (PSV) fusion during germination in Col-0 and *aspa* triple mutant seeds.

Figure 2-06. *ASPA2* expression during seed development.

Figure 2-07. *ASPA2* expression in vegetative and reproductive tissues.

Figure 2-08. Intracellular trafficking pathway of *ASPA2* in *Arabidopsis* roots.

Figure 2-09. Glycosylation and vacuolar trafficking of *ASPA2*-CFP.

Figure 2-10. Root architecture in Col-0 and *ASPA* mutants.

Figure 2-11. Transcriptional regulation of *ASPA1* and *ASPA2* and root growth in responses to ABA and NaCl treatments.

Figure 2-12. Propidium iodide (PI) staining in lateral root caps of Col-0 and *ASPA* mutants.

Figure 2-13. Fluorescent diacetate (FDA) and propidium iodide (PI) double staining in lateral root cap cells in Col and *ASPA* mutants over time.

Figure 3-01. Predicted structure of AtPSAPLIP1 and AtPSAPLIP2.

Figure 3-02. Predicted structure of saposin B (SapB)-like domains in AtPSAPLIP1 and AtPSAPLIP2.

Figure 3-03. Sequence alignment of plant PSAPLIPs from some angiosperms.

Figure 3-04. Primary structure similarity between plant PSAPLIPs with human prosaposin.

Figure 3-05. *AtPSAPLIP1* promoter GUS staining.

Figure 3-06. *AtSAPLIP2* promoter GUS staining.

Figure S01. Colocalization between autophagy marker ATG8a and *ASPA2*

Figure S02. Phenotype of 30 DAG plants of *ASPAs* overexpression.

Figure S03. Phenotype of 40 DAG plants of *ASPAs* overexpression.

Figure S04. 35S::*ASP2* D107A N404A-CFP (potential glycosylation site mutation) subcellular localization.

Figure S05. *ASP1* expression in seedlings.

Figure S06. *ASP3* promoter::YFP expression in lateral root cap.

Figure S07. Sequence alignment of PSAPLIPs in green algae, liverwort, moss and gymnosperm.

Figure S08. Sequence alignment of PSAPLIPs in angiosperms.

Figure S09. Sequence alignment of PSAPLIPs which contain three SapB-like domains.

Figure S10. Root growth in *AT3g51730* overexpression plants.

Figure S11. Root growth in *AT5g01800* overexpression plants.

Figure S12. Phenotype of 30 DAG Col and 35S::*AtPSAPLIP1*-CFP plants.

Figure S13. Phenotype of 30 DAG Col and 35S::*AtPSAPLIP2*-CFP plants.

Figure S14. *Arabidopsis* PSAPLIPs promoter :: GUS activity in seedlings.

Figure S15. Phenotype of *At5g01800* CRISPR mutant candidate.

Figure S16. Phenotype of *At3g51730* CRISPR mutant candidate.

Figure S17. Phylogenetic tree of PSAPLIPs in plants

Chapter 1: Literature review: Saposin-like proteins in plants

Abstract

The plasma membrane and endomembrane system is an essential component of all eukaryotic cells. Sphingolipids and microdomain-associated proteins that are associated with the plasma membrane and endomembrane system are important in plant growth and development. Elucidating functions of these proteins advances understanding of signal transduction from plasma membrane into cytosol and between different intracellular membrane compartments. Saposins and saposin-like proteins (SAPLIP) are among these proteins. SAPLIPs are a group of small proteins which usually consist of around eighty amino acids. Their main function is interacting with membranes. In plants, two types of proteins contain saposin B-like domains: aspartic proteases (ASPAs) and prosaposin-like proteins (PSAPLIPs). This review focuses on the functions of saposin-like proteins in animals, the reported saposin-like proteins in plants, and the knowledge gaps between plant saposin-like protein functions *in vitro* and *in vivo*.

Introduction: Sphingolipids and microdomains in plants

Sphingolipids are comprised of a lipid backbone and aromatic amino acid alcohol,

predominantly sphingosine, and there are at least 500 different molecular species of sphingolipids (Futerman et al., 2004). Sphingolipids are present in eukaryotes and some bacteria, and there are 168 different sphingolipids in *Arabidopsis thaliana* (Markham and Jaworski, 2007). In general, plant sphingolipids can be classified into four groups: glycosyl inositol phosphoceramides (GIPCs), glycosylceramides (GlcCers), ceramides, free long-chain bases (LCBs) (Pata et al., 2010). GIPCs are the predominant forms of complex sphingolipids in fungi and plants, but not found in animals (Warnecke & Heinz, 2003; Worrall et al., 2003; Lynch & Dunn, 2004). Some sphingolipids are only found in certain species or tissues. For example, long chain base (LCB) d18:2 containing GlcCer is the most abundant GlcCer in tomato and soybean plants. In these two species, GlcCer distributes throughout the plant (Markham et al., 2006), while GlcCer is mainly found in flowers and pollens in *Arabidopsis* (Michaelson et al., 2009). Further, sphingolipid species and levels can change during development, such as in olive fruit where it increases at during fruit development and then decrease upon fruit ripening (Ines et al., 2008). This high diversity of the molecule and the regulation of its biosynthesis signifies its versatile functions in plant physiology.

Sphingolipids are important components in the plant plasma membrane (PM)s and endomembrane system together with lipids, glycerolipids and sterols. In tobacco (*Nicotiana tabacum*) 'Bright Yellow 2' cells, GIPCs represent as high as 40% of the total PM lipids and 60% to 80% of total outer leaflet lipids (Cacas et al., 2016). Their structure contributes to the fluidity and biophysical properties of membranes (Huby

et al., 2019). For example, GlcCers have been implicated in chilling/freezing tolerance (Lynch and Steponkus, 1987; Imai et al., 1995; Minami et al., 2009; Takahashi et al., 2016). The *Arabidopsis thaliana* loss-of-function sphingolipid biosynthesis double mutant *sld1sld2* (sphingolipid Δ 8 long-chain base desaturases) is sensitive to cold (Chen et al., 2012). In addition to its role as a critical component in membranes, sphingolipids also show other biological functions. For example, sphingolipids are involved in programmed cell death (PCD) signaling transduction during plant development (Broderson et al., 2002; Liang et al., 2003) and immunity (Spassieva et al., 2002). Sphingolipids are also found to be necessary for sorting the membrane auxin carriers *AUX1* and *PIN1* from the trans-Golgi network (TGN) toward the plasma membrane (Markham et al., 2011) and glycosphingolipids with very long acyl chains stimulate lipid bilayer fusion during exocytosis and cytokinesis (Molino et al., 2014).

The various species of sphingolipids may determine the biophysical and biochemical bases for microdomains formation in membranes, and these membrane microdomains may be the structural bases for the various functions of sphingolipids as a result. The microdomain model comes from experimental results such as biochemical definition of detergent resistant membranes as well as live cell imaging (Lagerholm et al., 2005; Day et al., 2009). Microdomains are lipid-ordered domains enriched in sterols and sphingolipids. They exhibit self-assembly and recruit specific proteins into their regions (Yu et al., 2020). Some sphingolipids are enriched in microdomains, such as polyphosphoinositides (PI4P and PI4,5P2) (Furt et al., 2010)

and some structural phospholipids are rarely found such as phosphatidylcholines and phosphatidic acids (Mongrand et al., 2004; Laloi et al., 2007). These heterogeneous membranes provide different environments for lipid-protein interactions. Some proteins can move in and out of microdomains such as aquaporin *PIP2;1* (Li et al., 2011) and bacterial pathogen-associated molecular pattern flagellin (Flg22) (Keinath et al., 2010), while other proteins are very stable such as ATP Binding Cassette subfamily B (ABCB) transporters (Titapiwatanakun et al., 2008).

The microdomains can integrate signals by recruiting specific proteins into the microdomains. Some proteins are exclusively localized in microdomains, such as flotillins (Borner et al., 2005) and remorins (Konrad et al., 2014). Their function may be recruit other proteins into microdomains required for physiological processes. Microdomains are important for biotic stress for aggregation of Flg22. During abiotic stress such as drought stress, *AtFlot1* is involved in microdomain-mediated endocytosis of aquaporin *PIP2;1* in *Arabidopsis* (Li et al., 2011). Microdomains also directly affect plant growth and development. For example, the auxin transporters ABCB1 and ABCB19 are localized in microdomains and these two proteins stabilize *PIN1* auxin carrier in microdomains (Titapiwatanakun et al., 2008). In sphingolipid-defective mutants where microdomain structure is affected, ABCB19 is not properly trafficked or localized to the plasma membranes (Yang et al., 2012).

In general, sphingolipids and microdomains are important for numerous physiological pathways in plant cells. Many of these functions are executed by the

interacting or associated proteins. One of these proteins is called Sphingolipid Activator ProO[S]teIN (SAPOSIN). These proteins are involved in sphingolipid metabolism in human cells. But in other species, the functions of saposin-like proteins (SAPLIPs) appear to have additional functions.

Saposin-like Proteins (SAPLIPs)

Saposin-like proteins (SAPLIPs) are named after saposins, which are four small proteins (Saposin A through D) derived from one single precursor called prosaposin. Saposins are important in cellular metabolism as cofactors in sphingolipid catabolism in human cells (Bruhn, 2005).

SAPLIPs are found throughout eukaryotes from amoebozoans to mammals. They are not present in prokaryotes, except that three bacterial sequences have been assigned to this family in the InterPro database (<http://www.ebi.ac.uk/interpro>). However, these sequences lack the typical pattern of cysteine residues required to form the SAPLIP secondary and tertiary structures (accession number Q9FBA5 from *Borrelia hermsii*, accession number Q5XZA4 from *Borrelia garinii* and accession number Q5X236 from *Legionella pneumophila*) (Bruhn, 2005).

From phylogenetic data, sequence similarity among SAPLIPs is usually below 25%, which is the general threshold for a gene called a homolog (Bruhn, 2005). However, the data do show that SAPLIP domains evolved from an ancestral protein. Gene

duplication and subsequent mutations during evolution lead to functional versatility (Bruhn, 2005). The general features of a SAPLIP domain are six conservative cysteines and several conservative hydrophobic and polar charged residues (Bruhn, 2005). These cysteines form three pairs of disulfide bonds and together with the hydrophobic residues, forming a hydrophobic cave which allows lipid interaction.

Biological functions of SAPLIPs in animals

Human SAPLIPs are among the most well-studied SAPLIPs. From those studies, the function of SAPLIPs may be categorized into three types: (i) Membrane targeting (Figure 1-01A); (ii) membrane perturbation without lipid extraction (Figure 1-01B); (iii) membrane perturbation and lipid extraction (Figure 1-01C) (Bruhn, 2005).

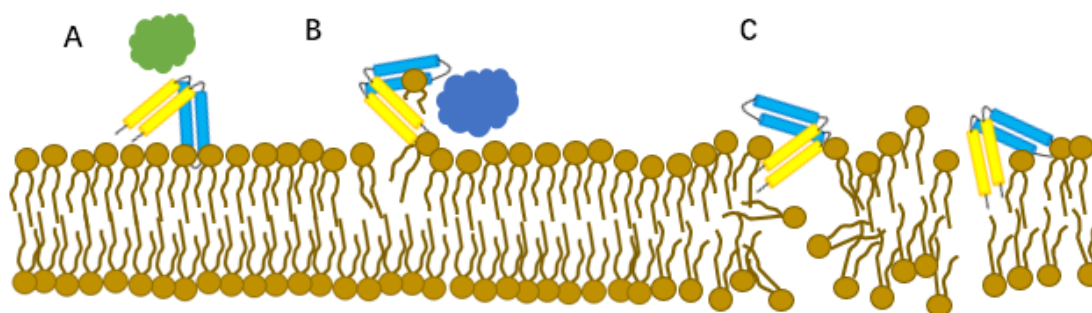


Figure 1-01. Schematic depiction of the three main activities of SAPLIPs. (A) Membrane targeting by the SAPLIP domain. (B) Presentation of lipids as substrate for an independent enzyme, either by extraction from the membrane or by disturbance of the well-packed lipid order. (C) Membrane permeabilization by perturbation owing to single molecules or by pore-formation of oligomeric proteins. Yellow and blue bars,

SAPLIP domain; green cloud, enzymatic domain; blue clouds, independent enzyme acting on lipids (arrows). Image is modified from Bruhn, 2005.

SAPLIP domains may exist as an independent functional unit or as a part of a multidomain protein. In animals, SAPLIPs are found to participate in a variety of different functions. For example, they are co-factors of lipid-degrading enzymes (Kishimoto et al., 1992; Schuette et al., 2001), surfactant tension regulator surfactant protein B (Cochrane et al., 1991), and the antimicrobial effector NK-lysin (Pena et al., 1997). Studies show that saposins can extract lipids from membranes and load them on to the antigen-presenting molecules Cluster of Differentiation 1d (CD1d) (Zhou et al., 2004; Winau et al., 2004; Kang et al., 2004). Saposins A, B and C are implicated in various disease states whereas no known deficiency corresponding to loss of saposin D in humans has been documented. However, a saposin D mouse knockout resulted in deleterious effects (Matsuda, 2008). In general, defective saposin-disease states arise from the accumulation of ceramide derivatives in various tissues resulting in pathological states. NK-lysin is a member of the saposin-like protein family and an antimicrobial and antitumor polypeptide. It also has lytic activities against bacteria, fungi and protozoan parasites (Hong et al., 2008).

Although most SAPLIP function is based on lipid binding property, recent studies found that some SAPLIP activities are independent of lipid interactions. One example is crystallin which functions in lens transparency in eyes. J3 crystallin, containing two

SAPLIP domains, is found in the transparent jellyfish *Tripedalia cystophora* (Piatigorsky et al., 1997). This raises the hypothesis that SAPLIPs are not only capable of lipid interactions, but also capable of protein-protein interactions in some cases.

The function of SAPLIP multidomain proteins is less studied compared to the autonomous units. One example of a multidomain SAPLIP is the human acyloxy acylase. After proteolytic processing of precursor, the SAPLIP domain and the catalytic domain appear to be linked by a disulfide bond. The SAPLIP domain appears to be required for intracellular targeting and catalytic activity of the acylase (Staab et al., 1994), and therefore, the SAPLIP domain may contribute to the enzyme activation.

The application of novel research methods will reveal more details to the function of SAPLIPs in multidomain proteins.

SAPLIPs in plants

SAPLIPs are found in from green algae to flowering plants. All reported plant SAPLIPs are characterized as domains or insertions in a subset of aspartic proteases. These insertions are often called plant specific inserts (PSI) because this is not found in animal aspartic proteases (Shown in Figure 1-02).

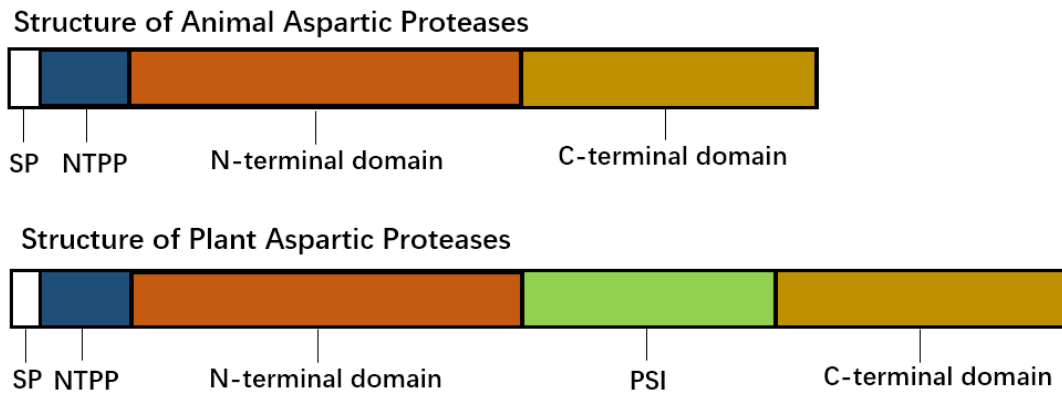


Figure 1-02. Schematic comparison between animal aspartic proteases and plant aspartic proteases that contain the saposin domain/plant specific insert. SP: signal peptide; NTPP: N-terminal propeptide; PSI: Plant specific insert.

Interestingly, SAPLIPs in plants are the model of circular permutation: the orientation of helices is switched from N terminus to C terminus. As a result, they are sometimes called “swaposins” (Blivem et al., 2012). The overall configuration of the secondary and tertiary structure is not affected. In addition to SAPLIPs contained in some aspartyl proteases, there are also independent SAPLIPs in plants. So far, there are no reports on either the structure or the biological functions of those independent plant SAPLIPs.

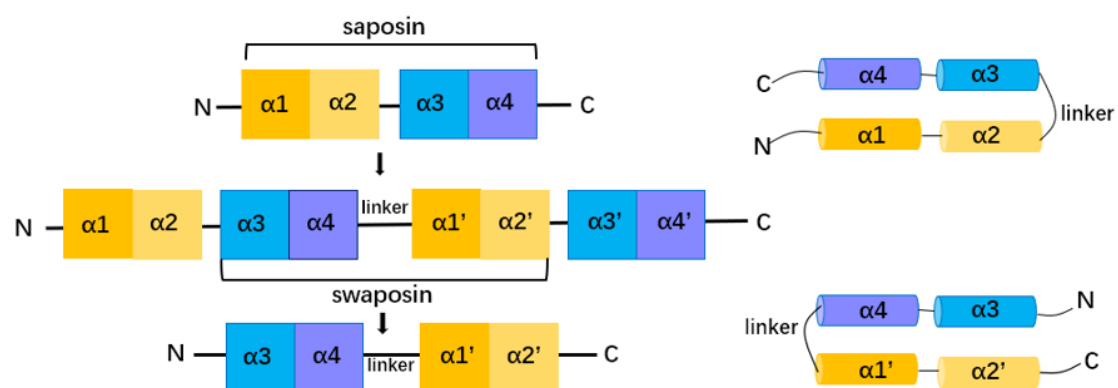


Figure 1-03. Schematic comparison between saposin and swaposin. The order of the helices in the swaposin is permuted relative to the saposin. However, the structure is conserved. Image is modified from Bliven et al. (2012).

General structure features of SAPLIPs

As mentioned above, SAPLIPs are highly diverse proteins with amino acid similarities below 30%. Although no common shared motif is found, there are conserved features: the distribution of hydrophobic amino acids forming the core and six conserved cysteine residues which form the disulfide bonds. Most reported SAPLIPs share a similar secondary structure (Bruhn, 2005).

Using the NK-lysin as a typical example, 5 helices fold into two halves. The first half consists of helices 4 and 5 packed perpendicularly against helix 1. The other half contains helix 2 and 3 (Liepinsh et al. 1997) (Figure 1-04A). Saposin B is representative of other types of SAPLIPs. The two halves of saposin B crystallizes as a dimer. The saposin B monomer shows an open formation in a V shape. This has been proposed as the lipid binding position (Ahn et al., 2003). Studies from human saposins would provide information about how it functions in the cell.

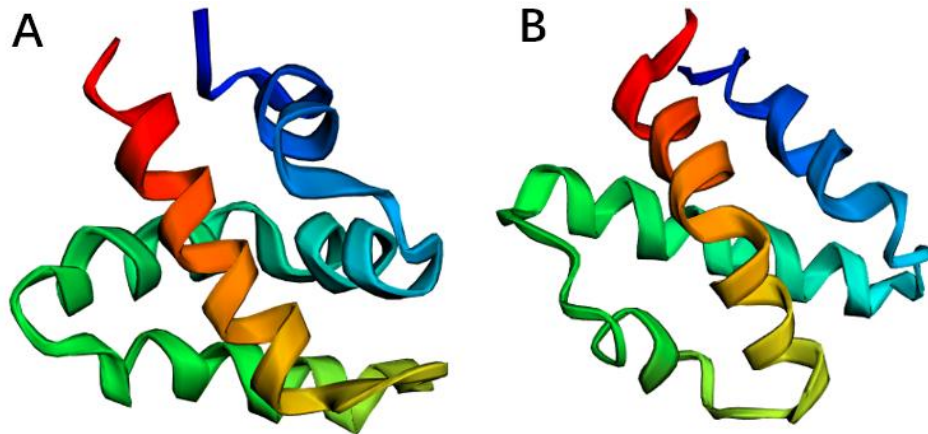


Figure 1-04. Schematic illustration of saposin domain of NK-lysin and the swaposin domain of prophytepsin from barley. (A) The saposin domain of NK-lysin (Protein Data Bank 1NKL). (B) the swaposin domain of prophytepsin (Protein Data Bank ID 1QDM). Sequences were downloaded from Protein Data Bank. Structures were constructed in Phyre2 and illustrated in EzMol.

Structure and function of human saposins.

Structure of and function of prosaposins

Saposins are processed from prosaposins. Prosaposins exist in three different forms - as a precursor for mature saposins, in a secreted form and as an integral membrane component (Hiraiwa et al., 1993). Prosaposin as a precursor for mature saposins is the well described but the latter two forms are not as well elucidated. Prosaposin precursor is found in two major forms, a 68 kDa intracellular form and a 73kDa extracellular form. Both forms are processed to a 50 kDa protein after

deglycosylation by N-glycosidase. In addition, prosaposin exists as a multimer at neutral pH and as a dimer in acidic pH (Hiraiwa et al., 1993).

The prosaposin is biosynthesized, glycosylated, secreted extracellularly and it is proteolytically processed in the intracellular space of the lysosome to generate mature saposin A, B, C and D (Kishimoto et al., 1992). The signal peptide (first 16 amino acid residue) is cleaved from the preprosaposin to generate the prosaposin, as demonstrated in both rat and human milk (Kishimoto et al., 1992). Prosaposins can stimulate lysosomal β -glucosidase and β -galactosidase activities bound to gangliosides, which is like mature saposins. However, prosaposins cannot stimulate hydrolysis of sulfatides, while mature saposin B have this activity (Hiraiwa et al., 1993). A thiol protease was shown to catalyze the proteolysis of the recombinant prosaposin to subsequent mature saposin domains (A, B, C and D) by cleavage at the peptide linkages (Kishimoto et al., 1992).

Two fragments, 39 kDa and 26 kDa, from partially purified samples of recombinant prosaposin cross-reacted with the anti-saposin C antibody. Through N-terminal sequencing, the 39 kDa protein was found to be produced by cleavage between leucine179 and phenylalanine180, corresponding to the linkage between saposin A and B, leaving a tri-saposin composed of B, C and D. The 26 kDa protein was produced by the cleavage between glutamic acid297 and leucine298 between saposins B and C, leaving a di-saposin of C and D (Hiraiwa et al., 1993; Kishimoto et al., 1992). In human cells, trisaposin B-C-D can also be processed into saposin D between

leucine 387 and cysteine 388, and produce the disaposin B-C. It is unable to distinguish between the pathway that liberates saposin D or saposin B from the trisaposin quantitatively (Leonova et al., 1996). The cleavage site is different for human seminal plasma prosaposin and insect prosaposin and that the cleavage of saposin A generates a derivative with 20 extra residues from the N-terminus. This suggests that post proteolysis activities are required to generate mature saposin A (Kishimoto et al., 1992). In insect cells, prosaposins are predominantly processed into A-B and C-D disaposins, and only small amount of mature saposins could be detected (Leonova et al., 1996).

These findings indicate that the mature saposins come from cleavage between saposin A and B, B and C, and C and D.

Structure and function of mature saposins

Mature saposins A, B, C and D are structurally similar and are composed of six cysteines forming three intramolecular disulfide bonds, a glycosylation site and conserved prolines in identical positions (Kishimoto et al., 1992).

Saposins are considered highly dense and firmly disulfide-linked molecules due to their high heat stability, extensive disulfide linkages and resistance to many proteases (O'Brien and Kishimoto, 1991).

In addition, saposins are glycoproteins with high levels of carbohydrates.

Approximately 40% of total glycosylation events are found in saposin A and about 20% were present in saposins B, C and D. Saposin A is also found to have two N-linked chains, whereas in saposins B, C and D, only one N-linked chain is present (Yamashita et al., 1990). However, the carbohydrate moiety is not essential for the activation of glucosylceramides (Sano and Radin, 1988).

Each saposin is composed of approximately 80 amino acid residues. Saposin A is between amino acids 60 and 143 in the prosaposin and activates β -glucosylceramidase, β -glucosidase and β -galactosylceramidase (Fabbro and Grabowski, 1991). Saposin B is between amino acids 195 and 275 in the prosaposin and activates arylsulfatase A, α -galactosidase A, GM1- β -galactosidase and various other enzymes. Saposin C is between amino acids 311 and 390 in the prosaposin and activates β -glucosylceramidase, β -glucosidase and β -galactosylceramidase. Lastly, saposin D between amino acids 405 and 487 in the prosaposin is responsible for the activation of sphingomyelinase (O'Brien and Kishimoto, 1991). Proteolytic cleavage is critical for the formation of these mature saposins.

The four saposins differ in their hinge regions and in the alpha helix-3 sites which may allow conformational changes during association with lipids or the lipid bilayer. ^{15}N labeled NMR spectroscopy of all four human saposins at both neutral and acidic pH showed that the mature saposins were highly unstable at pH 4.0 (John et al., 2006), but exhibited maximal α -helical stability at pH 4.5, the optimal pH for most lysosomal hydrolases. This finding suggested that their α -helical structures were important for

their physiological functions.

Structural features of saposin A

Under neutral conditions, saposin A is monomeric in a closed conformation. At lysosomal pH 4.8 saposin A forms a dimer, but remains in the closed conformation (Hill et al., 2015). Only in the presence of lipids or detergents does it undergo conformational change into the open state, forming lipo-protein particles with a variety of lipids (Ahn et al., 2006; Popovic et al., 2012; Hill et al., 2015).

The crystal structure of open saposin A dimer with detergent lauryldimethylamine-N-oxide (LDAO) shows that 40 LDAO molecules are enclosed in the two open chains. This suggests that the dimer configuration shields the hydrophobic surface sides of monomers (Ahn et al., 2006).

Structure feature of saposin C

Saposin C shows similar pH- and detergent-induced oligomerization (Ahn et al., 2006). Saposin C is monomeric under neutral conditions. It has been reported to be dimeric (John et al., 2006; Rossmann et al., 2008) or trimeric (Ahn et al., 2006) in solution at low pH. Saposin C remains the closed and compact conformation which shields hydrophobic residues in the absence of lipids. In the presence of SDS micelles, it changes to an open V-shaped conformation (Haukins et al., 2005).

Saposin C interaction with membranes might be facilitated by neutralization of acidic residues. Negatively charged surfaces might create electrostatic repulsion from negatively charged groups of membrane lipids (De Alba et al., 2003; Hawkins et al., 2005). However, about 50% of the glutamates were neutralized at pH 5 in saposin C by pH titration measurements, although no conformational change occurred between pH 5 and 7 (Hawkins et al., 2005). As a result, several lysines in saposin C are proposed to contribute to interactions with membranes (Hawkins et al., 2005).

Structure feature of saposin B

Saposin B is slightly different from saposin A and C. Saposin B has been reported as the primary saposin facilitating lipid binding to CD1d molecules (Yuan et al., 2007), although all saposins promote lipid binding to CD1d. The first 24 N-terminal amino acids residues of saposins B appear to form β -sheet configurations, while in saposins A, C and D, the helical structures are predominant (Chou and Fasman, 1978). Circular dichroism analysis has also shown that saposin B has high β -sheet content (O'Brien and Kishimoto (2001). Unlike saposin A and C whose dimerization requires the presence of lipid or detergents, saposin B dimerizes at neutral and low pH, either with or without detergents (Ahn et al., 2006; Popovic and Prive, 2008). The notable feature of the dimer is the V-shaped hydrophobic open cave formed by clasping monomers. The dimers may bind one or more lipid molecules where lipid polar headgroups

remains in the solvent (Ahn et al., 2003; Ciaffoni et al., 2006). The saposin B pH optimum is 6, which is higher than lysosomal pH. The affinity for phospholipid membranes of saposins A, C, and D depends on low pH, in contrast to saposin B. This suggests that saposin B might facilitate lipid binding to CD1d throughout the endomembrane system (Yuan et al., 2007). Saposin B may bind, transport and transfer a large variety of membrane sphingolipids and phospholipids to lysosomes (Ciaffoni et al., 2006). In general, Saposin B seems function as a lipid extractor and solubilizer that interacts transiently with membranes. Reports show that saposin B extracts target lipids from membranes and forms soluble protein-lipid complexes in open conformation dimeric state (Ahn et al., 2003).

Structure feature of saposin D

Saposin D is the least studied saposin compared to the other three. Saposin D is a ceramide activator protein involved in activation of hydrolysis of ceramide to fatty acids and sphingosines by acid ceramidases (Klein et al., 1994; Linke et al., 2001). Unlike to other saposins, saposin D (SapD) crystal structures show a compact closed monomeric form both as neutral and acid pH (Rossmann et al., 2008; Popovic et al., 2008). However, there is still the possibility that SapD forms dimers (Popovic et al., 2008). At low pH and in the presence of phospholipids, saposin D shows lipid binding activity and sphingolipid activation function (Ciaffoni et al., 2001; Linke et al., 2001;

Popovic et al., 2008).

Mechanistic model of saposin-lipid interactions

Rossmann et al. (2008) proposed a lipid solubilizer model for saposin D. Before the interactions with membranes, saposin D is in a monomer-dimer equilibrium in a closed, compact configuration. The low pH in the lysosomes neutralizes negatively charged glutamates and possible other residues, and thereby makes saposin D more hydrophobic and reduces repulsion of saposin D by negatively charged membrane surface. The “bottom” of saposin D, which contains the positively charged amino acids, likely interacts on the intralysosomal membranes that are enriched with negatively charged lipids (Figure 1-05A). On the other side, nonpolar residues are enriched on the “top” of the proposed saposin D dimer. This interaction may lead to moving towards hydrophobic environment by rolling the dimer by 180° around its long axis on the membrane surface. Then the hydrophobic residues are brought into membrane bilayer and the positively charged residues are exposed to the solvent. During the initial interaction with the membrane, monomer-monomer interactions in the dimer are possibly weakened by structural rearrangements in saposin D. The interaction of carbohydrate moieties with the gatekeeper amino acids in saposin D (Phe50, Phe4 and Tyr54) hide the hydrophobic interior. This initiates a hinge-bending and opening of saposin D allowing hydrophobic surface of the α helices to insert into the membrane. Thus, the membrane structure is perturbed.

Saposin C functions in vesicle fusion and destabilization after the initial binding to negatively charged membranes in a manner similar to saposin D before gate opening, membrane insertion and transition to an open configuration (Figure 1-05C). Helix pairs $\alpha 1/\alpha 4$ at both ends of saposin C dimers clip to opposing liposomal vesicles. As a result, the vesicles are brought close enough for fusion to occur (Rossmann et al., 2008). The schematic summary is illustrated in Figure1-06.

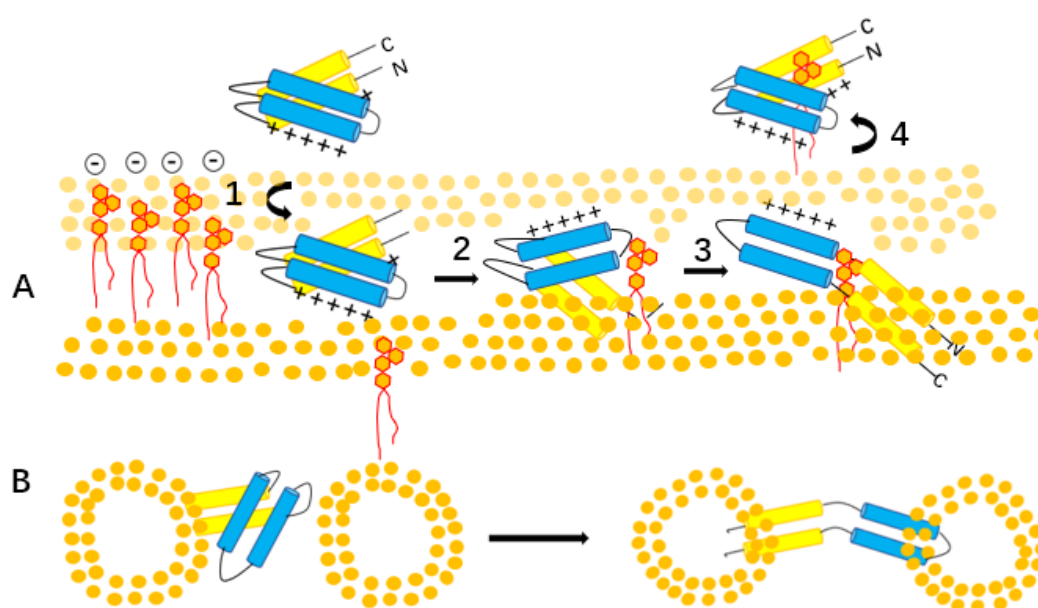


Figure 1-05. Models of lipid activation by saposins. (A) Schematic model for saposin D (SapD)-stimulated lipid activation. Step 1: water-soluble SapD monomers and dimers bind to the negatively charged membrane surface. Step 2: SapD rotates such that the hydrophobic “top” of the dimer faces the membrane surface. Step 3: SapD changes configuration into a boomerang shape also found in saposin C (SapC)(below), and amphipathic α helices stretch parallel to the lipid bilayer, exposing polar residues to the solvent. The hydrophobic surface of the saposin dips into the membrane and

perturbs its structure. Step 4: SapD changes configuration in the closed form, lifts a lipid molecule out of the membrane, and may leave the membrane with bound lipid.

(B) Clip-on model for SapC-induced vesicle fusion proposed by Wang et al., 2003. SapC molecules anchored to phospholipid bilayers of vesicles clip to adjacent membrane layer, bringing the vesicles close enough for fusion. The size of the vesicle and saposins are not on the same scale. Image and descriptions are adopted from Rossmann et al. (2008).

In summary, the general working model for SAPLIP family would be like the following: the soluble, monomeric form of SAPLIP holds a closed conformation with the hydrophobic surface hidden in the cavity. Charged residues mediated the initial contact with the negatively charged lipid membrane surface by electrostatic interactions. Then the protein change into open conformation. This change would lead to dimerization or oligomerization. This is speculated that a deeper perturbation of the membrane by interaction between the cavity and the lipid acyl chains. The membrane-embedded oligomer is hypothesized to form a pore in the membrane allowing presentation to the hydrolytic enzymes.

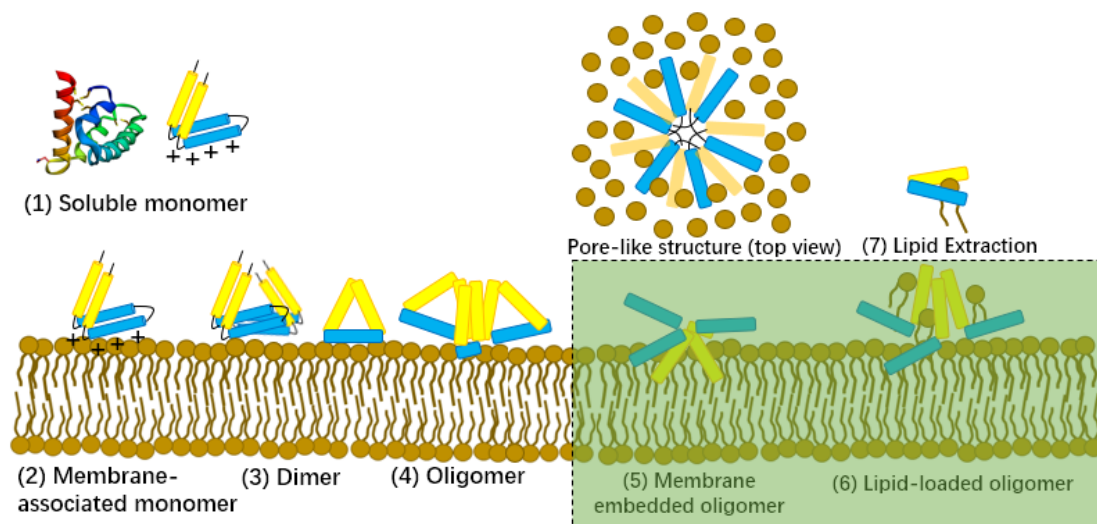


Figure 1-06. A model for the interaction with membranes and a general mechanism of action of the SAPLIP family. Figures inside the dotted green box are entirely speculative configuration that could be adopted by members of this family. (1) Soluble monomer in the solution (2) Monomer associates with membrane surface (3) Dimerization occurs on membrane surface (4) Dimers form into oligomers (5) helices insert into the membrane, and probably create a pore structure on membrane (top) (6) Saposin-like proteins loaded with lipid molecules (7) SAPLIP leaves membrane with lipid molecules. Image is adopted from Olmeda et al. (2012).

Structure of SAPLIPs in plants

The structure of PSIs in plants are also reported, such as cardosin A in cadoon *Cynara cardunculus* (Frazao et al., 1999; Egas et al., 2000), StAP in potato *Solanum tuberosum* (Bryska et al., 2011), phytepsin in barley *Hordeum vulgare* (Kervinen et al., 1999). Together with *Arabidopsis* aspartic proteases, these four proteins are often studied as models of SAPLIPs in plants (Figure 1-07A).

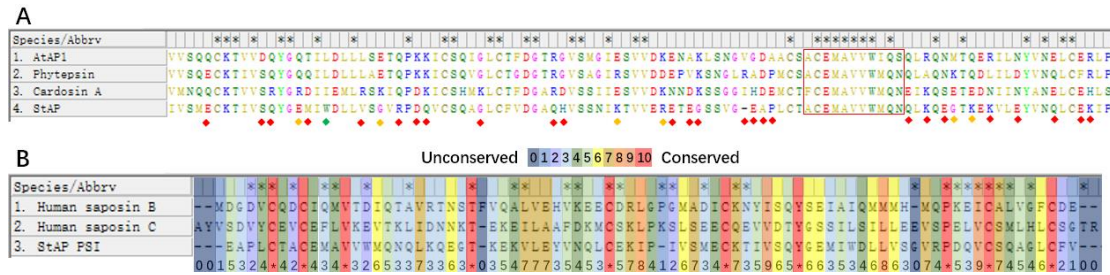


Figure 1-07. Plant PSI primary structure. Alignments with Clustal MUSCLE in Mega X.

(A) Multiple sequence alignments for PSI from AtAP (1), phytepsin (2), cardosin A (3) and StAP (4). Red diamonds: charge differences; orange diamonds: charge inversions; green diamond: additional Trp exclusive to StAP PSI. (B) Multiple sequence alignments for human saposin B (1), human saposin C (2) and StAP PSI (3) colored based on sequence conservation. StAP PSI is presented with its N-and C-terminal halves swapped to align with the saposins. Image and descriptions are adopted from Bryksa et al., 2017.

In general, plant PSIs are similar with human saposins in terms of sequence features and overall structure (Figure1-07B). The six conserved cysteines also appear in plant PSIs. PSI from phytepsin shows similarity with NK-lysin (Kervinen et al., 1999). PSI from cardoon shows high similarity to human saposins C and it can also activate human glucosylceramidase *in vitro* (Brodelius et al., 2005). PSI of StAP is also reported to show similarity to human saposin C and it also induces vesicle disruption *in vitro*. The secondary conformation is pH-dependent, which is similar to human saposins (Bryksa et al., 2011).

A recent study shows high sequence similarity and conservation among these

four PSI. They all show leakage activity in bilayer composed of a vacuole-like phospholipid mixture and membrane fusion activity *in vitro*. This activity is pH-dependent. The leakage activity is higher at pH 4.5 and requires the presence of acidic phospholipids such as phosphatidylserine. Low pH results in dimerization of potato PSI, and the monomer is prevalent under neutral pH. All the studies support that plant PSIs are similar to mammalian saposins in terms of structure and molecular activities.

As mentioned above, low pH activates bilayer membrane leakage activity. Conformation change is likely the molecular basis of this. A recent study found a novel 6-residue motif in H3 helix [N/Q]-[N/Q]-[N/Q]-[A/L/I/V]-[K/R]-[N/Q] which may contribute to this configuration change. A point mutation K83Q in this motif in helix H3 blocks the response to low pH activation with respect to conformation change (Bryksa et al., 2017). This motif may be responsible for lipid-interactions as this motif is also found in several other membrane-interacting proteins (Bryksa et al., 2017). This motif is not seen in human saposins.

As PSI is part of the aspartic protease, elucidating functions of aspartic proteases helps the better understanding the role of PSI in plant cells.

Plant aspartic proteases

General Information about plant aspartic protease

Proteases are an important for physiological processes and in commercial

applications. Proteases are one of the most important type of industrial use enzymes and they comprise approximately 60% of all commercial enzymes on the market (Feijoo-Siota and Villa, 2011). The diverse applications include food science and technology, the pharmaceutical industry, and detergent manufacturing. (Feijoo-Siota and Villa, 2011).

Aspartic proteases in *Arabidopsis*

In *Arabidopsis thaliana* genome, there are over 550 protease sequences categorized into five types: serine, cysteine, aspartic acid, metallo and threonine (MEROPS peptidase database, <http://merops.sanger.ac.uk/>). This reflects a wide variety of biological functions. (Beers et al., 2004).

Aspartic proteases (family A1) are characterized by a common bilobal tertiary structure containing two catalytic aspartic acid residues. They are found in all higher organisms. The most noticeable feature of A1 family is that there are two conservative aspartyl sites.

They are believed to be involved in the processing of propeptides in various plant tissues, such as in the breakdown of storage protein in seed germination (Belozersky et al. 1989) and the proteolytic processing and maturation of storage proteins (Hiraiwa et al. 1997b, Runeberg-Roos et al. 1994). They have also been shown to be involved in the turnover of pathogenesis-related proteins in tobaccos induced by stress (Rodrigo et

al. 1991), plant senescence and programmed cell death (Chen and Foolad 1997, Cordeiro et al. 1994).

Plant aspartic proteinases are also used by man for food processing. Protein extracts of *Cynara cardunculus* are used for cheese manufacturing (Cordeiro et al. 1992), and the aspartic proteinases from cocoa are important in the fermentation process of the beans for generation of flavor peptides from storage proteins (Biehl et al. 1985).

There are 59 annotated *Arabidopsis* A1 proteases identified. Predicted aspartic proteases from other families include the A11 family (approximately 45 members) retrotransposon endopeptidases and two presenilin-like proteins from A22 family (AAL24266 and AAD23630) (Beers et al., 2004). According to the sequence similarity, they are divided to five subfamilies, A1-1 (35 members), A1-2 (17 members), A1-3 (2 members), A1-4 (3 members) and A1-5 (2 members). The main difference between each subfamily is the number and distribution of exons and introns (Beers et al., 2004).

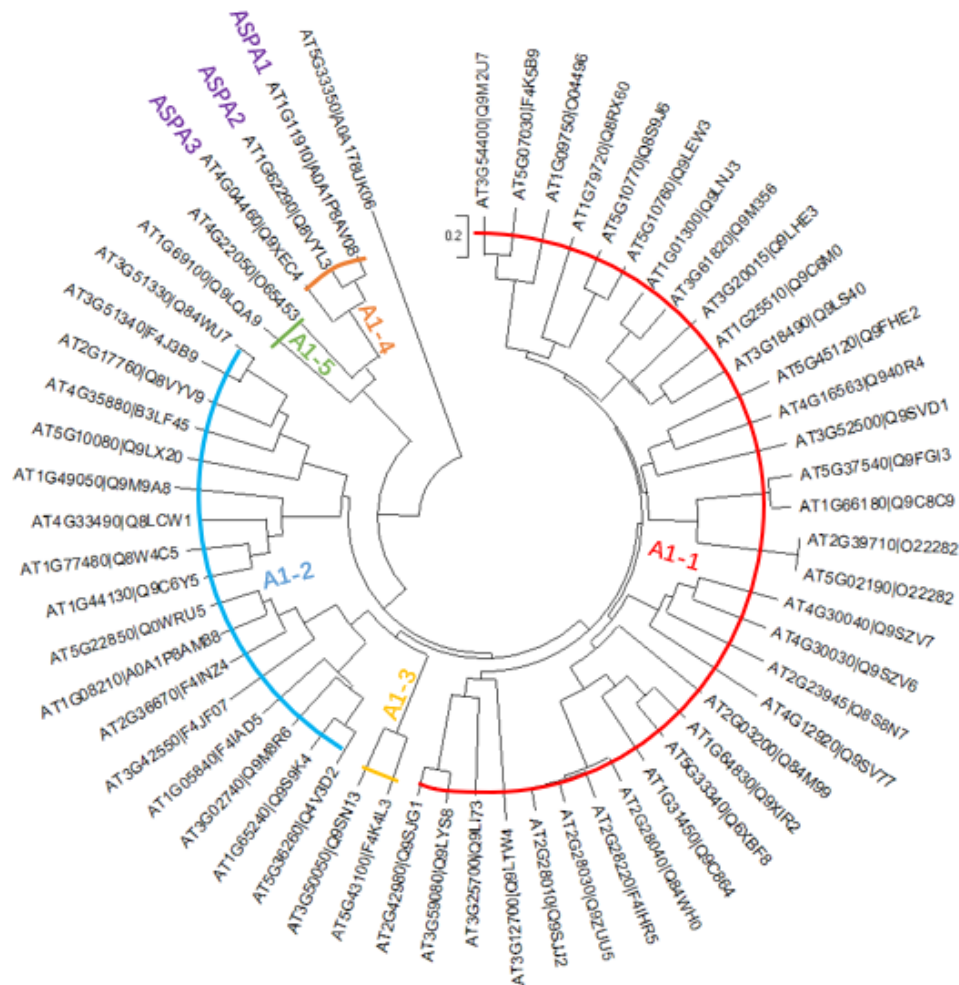


Figure 1-08. Phylogenetic tree of A1 family proteases in *Arabidopsis*. Neighbor-joining tree based on point accepted mutation (PAM) distances generated from an alignment of protein sequences in ClustalW. Five putative groups are indicated with colors. At5G33350 is mentioned in Beers et al., 2004 as a member in A1 family, but it doesn't contain the first conservative aspartic site. It is outside the five subgroups in this tree. *ASPA1* (At1g11910), *ASPA2* (At1g62290) and *ASPA3* (At4g04460) are indicated in purple characters. The subfamilies are based on Beers et al., 2004.

Several members in A1-1 subfamily are annotated as "chloroplast nucleoid DNA

binding protein-like” (CND41-like), originally identified from tobacco *CND41*(T01996). CND41 exhibits both proteolytic (Murakami et al., 2000) and chloroplast DNA-binding (Nakano et al., 1997) capabilities *in vitro*. No biological functions are reported for A1-3 and A1-5 subfamilies. A1-4 subfamily is the only one that contains the SAPLIP domain.

Biological functions of aspartic proteases in plants

The biological function of the A1-4 subfamily in several plant species have been reported. Cardosin A and cardosin B accumulate during seed maturation, and cardosin A is synthesized *de novo* at the time of radicle emergence. This suggests that cardosins are involved in storage protein processing during seed development and protein degradation during seed germination (Pereira et al., 2008). Cardosin B gene expression was also observed in pistils and ovules and is proposed to be involved in programmed cell death dependent degeneration of nucellus in cardoon (Figueiredo et al., 2006; Pereira et al., 2008). An aspartic protease in leaves of common bean (*Phaseolus vulgaris*) was found in a screen for drought tolerance (Cruz de Carvalho et al., 2001), and is upregulated by water stress in beans. A typical aspartic protease in pineapple fruit *Ananas comosus* was found to be upregulated by chilling treatment, which suggests a role in chilling stress resistance (Raimbault et al., 2013). Overexpression of a sweet potato aspartic protease SpAP1 promotes ethephon-induced leaf senescence (Chen et al., 2015). In the pitcher plant *Nepenthes alata*, of the five aspartic proteases

identified and four of them contained the SAPLIP domain. *NaAP2* and *NaAP4* transcripts were detected in the digestive glands which suggests that they are associated with secretion (An et al., 2002).

Arabidopsis has three typical aspartic proteases and they are also called phytepsins due to their similarity to mammalian enzymes pepsin and cathepsin D. The homolog in yeast *Saccharomyces cerevisiae* PEP4 is required for activation of several vacuolar zymogens (Van den Hazel et al., 1992). The biological functions of the A1-4 subfamily in *Arabidopsis* is not well studied, but the expression patterns may provide information about their biological functions.

There are three members of A1-4 subfamily in *Arabidopsis*, named *ASPA1* (At1g11910), *ASPA2* (At1g62290) and *ASPA3* (At4g04460) respectively. *ASPA1* mRNA is detected in all tissues and is abundant in leaves during daytime. *ASPA3* is primarily expressed in flowers and *ASPA2* is primarily expressed in seeds. (Chen et al., 2002). All three genes can be detected in developing siliques and seeds, which suggests their roles in seed development. Their expression patterns also suggest that they have multiple roles in *Arabidopsis* development.

All three proteases are targeted in the vacuoles (Otegui et al., 2006; Figure 2-09; Figure S07). In plants, there are two kinds of vacuoles: the central lytic vacuole and the protein storage vacuole. The central vacuole resembles animal and yeast vacuoles, while the protein storage vacuole is plant specific. Protein storage vacuoles are found in germinating seedlings (Paris et al. 1996; Swanson et al. 1998), nutrition storage

tissues such as tubers, leaves and tree bark (Müntz and Muntz 1998; Zouhar et al. 2010). The structure of protein storage vacuoles is complex: there is a “globoid cavity” exhibits an acidic environment similar to the central vacuole, and it is partitioned within a more neutral protein storage vacuole lumen (Jiang et al. 2001; Tse et al. 2007). The neutral lumen allows the storage proteins to stay for longer time from degradation. As a result, protein storage vacuoles have dual functions (Xiang et al. 2013). The intracellular trafficking pathway is similar between the lytic vacuole destination and the protein storage vacuole destination, but there are specific receptors in protein storage vacuoles (Hinz et al. 1999). Protein storage vacuoles reserve nitrogen in the form of storage proteins during seed maturation. Seed storage proteins are the major component of many agriculturally crops, such as legume seeds (40% dry weight) (Bradford and Bowley 2003; Atta et al. 2004; Gottschalk and Muller 2012). During germination, protein storage vacuoles fuse to form a single central vacuole, and seed storage proteins are catabolized to provide amino acids for protein biosynthesis (Jiang et al., 2001). The primary storage proteins found in mature seeds are classified based on solubility as albumins (water-soluble), globulins (salt-soluble), prolamins (alcohol-soluble) or glutelins (weak-acid/weak-base soluble) (Ferreira et al. 1999). In *Arabidopsis*, the predominant seed storage proteins are the 12S legumin-type globulins and the 2S napin-type albumins (Gruis et al. 2002). Seed storage proteins need to be post-translationally modified for stable and dense package. Both the 2S and 12S proteins are translated as long precursors, inserted into the ER lumen, and

undergo post-translational cleavage *en route* to the protein storage vacuoles (Paris et al. 1996; Hara-Nishimura et al. 1998a; Swanson et al. 1998; Gruis et al. 2002; Otegui et al. 2006). For the 12S globulin precursors, the ER signal peptide is cleaved after insertion into the ER lumen, and disulfide bonds form and the proteins assemble into trimers (Muntz, 1998), then transported via the Golgi body to the protein storage vacuoles. In the multivesicular body or prevacuolar compartment, 12S are processed by the enzymes known as Vacuolar Processing Enzymes (VPEs) which produce mature disulfide-linked α - and β - chains (Otegui et al., 2006; Baud et al., 2008). The VPEs are a family of asparagine-specific cysteine endopeptidases which cleave seed storage proteins *in vitro* and *in vivo* (Hara-Nishimura et al. 1991; Gruis et al 2002; Gruis et al. 2004). These proteases were identified in the seed, the leaf and the root. They are involved in senescence, programmed cell death and biotic defences (Hara-Nishimura et al. 1991; Kinoshita et al. 1995; Misas-Villamil 2013). The expression pattern of ASPA1 is in parallel with VPEs and it is believed that it is also involved in seed storage protein procession. Reports have shown that ASPA1 was highly expressed during embryo development, accumulated in protein storage vacuoles, and has been shown to cleave 2S seed storage protein napins *in vitro* (D'Hondt et al. 1993a; Mutlu et al. 1999; Otegui et al. 2006).

The involvement of VPEs in seed germination is not documented, but the detection of ASPA in seed germination is reported (Pereira et al., 2008). It is possible that ASPAs are involved in seed germination regulation and possibly degrade seed

storage proteins during germination.

Among the three ASPAs in *Arabidopsis*, the most studied gene is *ASPA3* due to its specific expression pattern. The *ASPA3* promoter-reporter constructs showed signals in almost all tissues that undergo programmed cell death (PCD), such as lateral root caps, tracheary elements in proxylem, fading petals, tapetum in stamens and endosperm in developing seeds (Fendrych et al., 2014; Olvera-Carrillo et al., 2015).

In plants, PCD can be categorized into two types: developmental PCD or environmental PCD (Daneva et al., 2016). Developmental PCD occurs in specific cell types such as tracheary elements in xylem, lateral root caps and tapetum in stamens, in order to facilitate normal growth and development. Developmental PCD can also be triggered conditionally by cell signaling, which could be seen in self-incompatibility responses (Wilkins et al. 2014; Petrov et al. 2015). Developmental PCD also occurs in all types of aging cells in the end of plant senescence (Klimešová et al. 2015). Environmental PCD occurs in response to stresses such as irradiation or pathogens (Wu et al. 2014). A portion of cells sacrifice in order to protect the remaining tissues. Comparative studies showed that these two types are different in transcriptional signaling (Olvera-Carrillo et al. 2015). However, the executing components downstream of many PCD-related transcriptional factors are shared in different cell types (Olvera-Carrillo et al, 2015; Huysmans et al., 2018), such as *ASPA3*, *BIFUNCTIONAL NUCLEASE1 (BFN1)*, *RIBONUCLEASE3 (RNS3)*, *CYSTEIN ENDOPEPTIDASE1 (CEP1)*, *DOMAIN OF UNKNOWN FUNCTION679 MEMBRANE*

PROTEIN4 (DMP4) (Olvera-Carrillo et al., 2015; Ye et al., 2020).

In terms of morphology, PCD can be divided into three types: apoptosis, autophagic cell death, and necrosis (Lockshin and Zakeri, 2004; Bras et al., 2005). In necrosis, organelles swell up and the plasma membrane ruptures to release the components. This type is less studied and is believed less controlled by genetic programming. In apoptosis, the cells shrink, DNA is fragmented into small pieces and cell component is compacted into small vesicles. In autophagic type, autophagic vacuoles are formed for degradation of cell components, but the cell doesn't necessarily die, which distinguish from apoptosis (Theresa et al., 2008).

The early events of apoptosis process include caspase signaling (Danon et al., 2004) and proteases synthesis in preparation for protein degradation and nucleases synthesis for DNA fragmentation (Fendrych et al., 2014). Take PCD in lateral root caps as an example, the events during PCD include a decrease in cytoplasmic pH, plasma membrane permeabilization, vacuolar collapse, and final degradation cell materials (Fendrych et al. 2014). *ASPA3* is believed to one of the proteases in this process, although the single mutant of *ASPA3* doesn't show a PCD-related phenotype in lateral root caps (Fendrych et al., 2014).

A recent study showed a potential role of *ASPA1* in drought tolerance by overexpression in *Arabidopsis* (Sebastián et al., 2020). Overexpression lines of *ASPA1* had longer primary root length under drought conditions. The overexpressors also had reduced stomata index, reduced stomata density and a smaller stomatic aperture

compared to wild type plants. Higher expression levels of genes related to ABA signaling and biosynthesis were also found in *ASPA1* overexpression lines. *ASPA1* promoter-GUS activity showed that *ASPA1* was induced by ABA in leaves.

These results indicate multiple roles of ASPAs in plant growth and development.

Structure of plant aspartic protease

Most known aspartic proteases are from a single chain proenzyme which are then proteolytically processed and then form either a homomeric or heterodimeric mature enzyme (Laloi et al., 2002; Simoes and Faro, 2004). The proenzyme is characterized by a hydrophobic signal peptide, an N-terminal propeptide of approximately 40 amino acid residues, and the mature protein region composed of an N-terminal domain and a C-terminal domain, separated by the plant specific insert (PSI, SAPLIP) of approximately 100 amino acids (Koelsch et al., 1994; Asakura et al., 1995).

Phytpsin undergoes several proteolytic cleavage steps to produce the two-chain form of the mature enzyme (Figure 1-09). Early processing involves insertional removal of the signal peptide at the endoplasmic reticulum (ER), followed by removal of the N-terminal propeptide (Glathe et al. 1998), and then PSI cleavage, followed/accompanied by cleavage 5 kDa upstream of the PSI cleavage point (Ala-378) gives a heavy chain and a light chain. This process is BFA sensitive, indicating that it occurs after the peptide has passed the Golgi. Finally, these chains are processed to mature forms by removal of the remaining PSI; this occurs only 24h after synthesis *in*

vivo (Glathe et al. 1998). This step is critical for activation of the enzyme. Proteolytic processing of procardosin A to its mature form is highly pH sensitive *in vitro*, with optimal processing at pH 4, but retarded processing at pH 3 or 5 (Castanheira et al. 2005). This processing is unlikely to be totally autocatalytic; although processing was active *in vitro* and inactivated by pepstatin A and cleavage patterns were different from those observed *in vivo* (Glathe et al. 1998). One interesting example is found in an aspartic protease cirsin from *Cirsium vulgare*. The procirsin was expressed in *Escherichia coli* and shown to be active without autocatalytically cleaving its propeptide domain (Lufrano et al., 2012). This contrasts with the acid-triggered autoactivation by pro-segment removal. Recombinant procirsin displayed all typical proteolytic features of aspartic proteinases such as optimum acidic pH, inhibition by pepstatin, cleavage between hydrophobic amino acids and strict dependence on two catalytic Asp residues for activity.

In general, complete activation of typical aspartic proteases requires correct processing at a pH conducive to efficient cleavage. The cDNA encoding the precursor of *AtASPA1* was expressed as a functional protein using the yeast *Pichia pastoris*. The mature form of the recombinant *AtASPA1* was found to be a heterodimeric glycosylated protein with a molecular mass of 47 kDa consisting of heavy and light chain components, approximate 32 and 16 kDa, respectively, linked by disulfide bonds. Glycosylation occurred via the plant specific insert in the light chain. The catalytic properties of the recombinant *AtASPA1* were similar to other plant aspartic

proteinases with activity in acid pH range, maximal activity at pH 4.0, K_m of 44 μM , and k_{cat} of 55 s^{-1} using a synthetic substrate. The enzyme was inhibited by pepstatin A (Miguel et al., 2008).

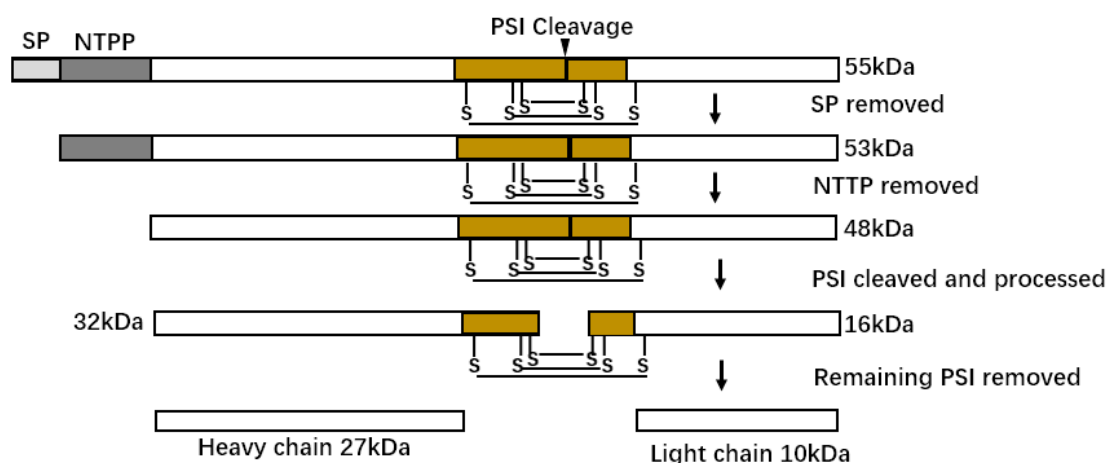


Figure 1-09. Schematic illustration of the protein processing steps for ASPA1. NTPP: N-terminal propeptide; SP: signal peptide; PSI: plant specific insert. First, the signal peptide (SP) is removed, then N-terminal propeptide (NTPP) is trimmed from the propeptide. Then the plant specific insert (PSI) is cleaved after transporting from Golgi body. The heavy chain (27kDa) and the light chain (10kDa) assemble into the mature protease. Destination of PSI is still unclear. Cleavage sites are predicted from alignment with cardosin A. Image is modified from Miguel et al., 2008.

Biological function of plant specific insert

The plant specific insert (PSI) distinguishes plant aspartic proteases from animal ones, and the biological functions of PSI have been a subject of study for a long time.

Studies show that PSI is likely to be critical for vacuolar targeting (Kervinen et

al.,1999; Terauchi et al., 2006). The intact proteases are targeted to vacuoles across the plant kingdom, from moss to seed plants (Schaaf et al., 2004; Kervinen et al., 1999; Terauchi et al., 2006). The deletion of PSI in soyAP2 resulted in the retention of peptides in ER (Terauchi et al., 2006).

The PSI brings phytepsin into contact with membranes, possibly with membrane-binding receptors proteins in Golgi apparatus. The prophytepsin is then trafficked to vacuoles and activated by proteolytic cleavage and the PSI is subsequently removed which breaks the interaction of the membrane receptor or the membrane itself (Kervinen et al., 1999).

The PSI also influences the route that phytepsin takes after leaving the endoplasmic reticulum, in addition to the vacuolar sorting function. It is currently unclear whether the PSI has two sorting signals: one for endoplasmic reticulum export or another one for vacuolar sorting or if the vacuolar sorting motif is recognized at the endoplasmic reticulum export site (Tormakangas et al., 2001).

Studies show that removing the plant specific domain has no effect on phytepsin activity; however, it does cause an accumulation of phytepsin in the extracellular space of the plant (Tormakagas et al., 2001). These findings support the role in vacuolar target and sorting, and the default secretion pathway proceeds without the PSI.

Another possible function of PSI is in pathogen defense pathway. *Solanum tuberosum* aspartic protease (StAP) PSI *in vitro* is able to kill spores of two potato pathogens in a dose-dependent manner without any deleterious effect on plant cells

(Muñoz et al., 2010). The StAP-PSI ability to kill microbial pathogens is dependent on the direct interaction of the protein with the microbial cell wall/or membrane, leading to increased fungi or bacteria cell permeability and lysis. StAP-PSI is able to kill human pathogenic bacteria in a dose dependent manner as well, but it is not toxic to human red blood cells at the concentrations and times assayed. Minimal bactericidal concentration (MBC) values determined for StAPs and StAP PSI are in the same order of magnitude as those previously reported for NK-lysin and granulysin (Muñoz et al., 2010). The constitutive expression of StAP-PSI induces defense genes in *Arabidopsis* and enhances *Arabidopsis* resistance against *Botrytis cinerea* infection. The StAP-PSI domain exerts cytotoxic activity toward *Botrytis cinerea*, and the constitutive expression of StAP-PSI increases growth in *A. thaliana* (Frey et al., 2018).

In a similar study, Pagano et al. (2007) investigated the importance of glycosylation, to the *Solanum tuberosum* aspartic protease. It was observed that aspartic protease accumulation into the apoplast of tubers and leaves after wounding required glycosylation (Pagano et al., 2007). This suggests that glycosylation may be necessary for *Solanum tuberosum* aspartic protease membrane and/or protein interactions. In a more recent study, the role of glycosylation was thought to affects intracellular trafficking of the aspartic protease.

The recombinant cardosin B PSI undergoes the conventional route from ER, the Golgi and the prevacuolar compartment to the vacuole if glycosylated. The non-glycosylated proteins entered vacuoles directly from ER and bypassed the Golgi bodies

(Vieira et al., 2019). This study suggests that there are unconventional trafficking pathways in the plant cell. It also elicits a more complex question: what's the exact function of glycosylation in PSI?

In general, most of reported data was based on the hypothesis that PSI is released during aspartic protease procession and secreted to the extracellular space. This needs more experimental data to support.

Perspective

Since the amino acid sequences of plant aspartic proteinases were described in the 1990s, a wide range of studies have explored their function in development largely and defense responses. *In vitro* biochemical and biophysical methods have been used to elucidate the function of the plant specific insert, but corresponding *in vivo* studies are largely lacking. There are some studies on the expression pattern of aspartic proteases, but there are few *in vivo* studies, especially the genetic studies. So far, no mutants of these aspartic proteases with growth and development phenotypes have been reported yet. A better understanding of these unique proteases could eventually be applied to agricultural production as a tool to manipulate plant growth and development.

For SAPLIPs in aspartic proteases, most studies are *in vitro*, and the *in vivo* studies are not with the whole length of the protein in context. This leads to a question: does

the PSI function independent from the protease or does it function in coordination with the protease? Or does it simply function as a signal for trafficking towards vacuoles? And what about the SAPLIPs which are independent units in plants? These SAPLIPs are called prosaposin-like proteins (PSAPLIPs) in this dissertation, and there are no published articles regarding these proteins to date. The biological functions of these PSAPLIPs also need to be explored. This will contribute to the current knowledge of plant cellular biology. With the *in vitro* studies of PSI and mammalian saposins, and the relatively small size of this protein family, this will also produce biological tools for cell biology in both scientific and industrial areas.

As a result, three questions are raised and will be discussed in this dissertation:

(1) What's the biological function of the typical aspartic protease and its PSI in *Arabidopsis*? Is there an independent role of PSI from the proteolytic domains? (2) How these aspartic proteases function in *Arabidopsis* cells? Are they involved in programmed cell death as proposed? (3) What's the feature of plant prosaposin-like proteins? What are the biological functions of the plant prosaposin-like proteins in growth and development?

Chapter 2 Functional study of saposin-like domain containing aspartic proteases in *Arabidopsis thaliana*

Abstract

Aspartic proteases (ASPAs) are important in plant growth and development. They are also one of the most important commercial proteases. Aspartic proteases containing a plant specific insert (PSI) sequence have been the subject of numerous studies. Several studies have reported the properties of PSI *in vitro*. However, few have reported the function of PSI or aspartic proteases that contain them *in vivo*. Here molecular genetic analysis has revealed that ASPAs were involved in seed maturation and regulate seed germination, root morphology in response to nitrogen supply, and were involved in programmed cell death (PCD) execution in *Arabidopsis*. Triple mutant *aspa1-2 aspa2-1 aspa3-3* showed delayed seed germination. Protein storage vacuoles fusing into the central vacuole was also delayed in the triple mutant during seed germination. *ASPA2* was expressed throughout the plant similar with *ASPA1*. *ASPA2* expression was responsive to environment stresses while *ASPA1* expression was stable. *ASPA2* was first imported into the endoplasmic reticulum and then through the endomembrane system: *trans*-Golgi network (TGN) then multivesicular bodies (MVB), and finally was transported to vacuoles. PSI was important for vacuolar localization, but site-directed mutagenesis revealed that the lipid binding motif inside

the PSI didn't seem to be required. ASPA2 colocalized with early endosome (EE) marker RabA3, which indicates the possibility for plasma membrane protein degradation. Root morphology was affected in *aspa* triple mutant. Primary root length was longer for triple mutant under low nitrogen levels. This is likely to result from the delayed programmed cell death of tracheary elements in the mutant xylem. Further analysis revealed that membrane permeability increased more slowly in the triple mutant lateral root cap cells during PCD. These results indicate that ASPA2 is involved in membrane disturbance during programmed cell death and modulate the rate of membrane permeability increase and therefore the rate of programmed cell death. These results indicate that ASPAs function in proteolytic activities in bulk and membrane disturbance. The independent role of PSI was not supported in this dissertation. This is the first time showing that ASPAs process seed storage *in vivo* and have impacts on seed maturation and seed germination. This is also the first time showing that ASPAs are involved in programmed cell death by promoting membrane permeability.

Introduction

Aspartic proteases are important in commercial application such as milk clotting for cheese making (Heimgartner et al., 1990). Aspartic proteases are also important in plant growth regulation. For example, cardosin A in cardoons is involved in storage protein processing during seed development and protein degradation during seed

germination (Figueiredo et al., 2006; Pereira et al., 2008). Aspartic proteases are also believed to be associated with drought tolerance in the common bean (*Phaseolus vulgaris*), chilling response in pineapple (*Ananas comosus*) and senescence in sweet potato leaves (Cruz de Carvalho et al., 2001; Raimbault et al., 2013; Chen et al., 2015). Aspartic proteases are also found in digestive ligands in pitcher plants (*Nepenthes alata*) (An et al., 2002).

The typical ASPA aspartic proteases contain an N-terminal propeptide and a plant specific insert (PSI) inside the protease peptide. Animal aspartic proteases lack the PSI sequence. The PSI resembles human saposin proteins in primary and secondary structures. Saposin-like proteins in animals have been well-studied and their major function is interacting with membrane lipids (Bruhn, 2005). In plants, the activity of the plant specific insert from the aspartic proteases have also been studied. This plant specific insert is critical for vacuolar targeting of the protease (Kervinen et al., 1999; Terauchi et al., 2006). The intact proteases are trafficked to vacuoles in the moss and soybean (Schaaf et al., 2004; Kervinen et al., 1999; Terauchi et al., 2006), and deletion of PSI in soyAP2 result in the retention of peptides in the endoplasmic reticulum (ER) (Terauchi et al., 2006). Studies also show that removing the plant specific insert has no effect on phytepsin activity; however, it does cause an accumulation of phytepsin in the extracellular space of the plant (Tormakagias et al., 2001). Other possible function of plant specific insert is in pathogen defense pathway. *Solanum tuberosum* aspartic protease (StAP) PSI *in vitro* is able to kill spores of two potato pathogens in a dose-

dependent manner without any deleterious effect on plant cells (Muñoz et al., 2010). The constitutive expression of StAP-PSI induces defense genes expression in *Arabidopsis* and enhances *Arabidopsis* resistance against *Botrytis cinerea* infection (Frey et al., 2018). The StAP-PSI domain exerts cytotoxic activity toward *Botrytis cinerea* (Frey et al., 2018). These findings suggest a role of PSI in plant defense, and it is assumed that this plant specific insert functions independently from the protease. A recent study identified a six amino acid motif in plant specific insert which accounts for conformation change and lipin bilayer fusion activity *in vitro* (Bryska et al., 2017). This motif is unique to plant saposin-like domains, which indicates that plant specific insert functions in a way different from animal saposin-like proteins. In plant cells, ASPAs undergo several proteolytic cleavage steps to produce the two-chain form of mature enzyme during which the PSI is released (Glathe et al., 1998). The destination of the released PSI is still not clear. This leads to the hypothesis that PSI function independently *in vivo* for normal plant growth and development.

In *Arabidopsis*, there are 59 annotated *Arabidopsis* A1 aspartic proteases identified (Beers et al., 2004). Only the A1-4 subfamily contains the saposin-like domain or plant specific insert. A1-4 subfamily has 3 members named *ASPA1* (At1g11910), *ASPA2* (At1g62290) and *ASPA3* (At4g04460), and they are also called phytepsins due to their similarity to mammalian enzymes pepsin and cathepsin D. The biological functions of the A1-4 subfamily in *Arabidopsis* is not well studied. Their homolog in yeast *Saccharomyces cerevisiae* PEP4 is reported for activation of several

vacuolar zymogens (Van den Hazel et al., 1992). This suggests that *Arabidopsis* ASPAs may also be activatice in the vacuoles.

The expression patterns may provide information about their biological functions. *ASPA1* mRNA is detected in all tissues. *ASPA3* is primarily expressed in flowers and *ASPA2* is primarily expressed in seeds. (Chen et al., 2002). These expression patterns suggest that they have multiple roles in *Arabidopsis* development.

ASPA1 is a well-known marker for prevacuolar compartment in developing seeds (Otegui et al., 2006) and believed to take part in seed storage proteins processing. In *Arabidopsis*, the predominant seed storage proteins are the 12S legumin-type globulins and the 2S napin-type albumins (Gruis et al. 2002). Seed storage proteins need to be post-translationally modified for stable and dense package. Both the 2S and 12S proteins are translated as long precursors, inserted into the ER lumen, and undergo post-translational cleavage *en route* to the protein storage vacuoles (Paris et al. 1996; Hara-Nishimura et al. 1998a; Swanson et al. 1998; Gruis et al. 2002; Otegui et al. 2006). In the prevacuolar compartment, seed storage proteins are processed by the enzymes known as Vacuolar Processing Enzymes (VPEs) which produce mature disulfide-linked α - and β - chains (Otegui et al., 2006; Baud et al., 2008). The VPEs are a family of asparagine-specific cysteine endopeptidases which cleave seed storage proteins *in vitro* and *in vivo* (Hara-Nishimura et al. 1991; Gruis et al 2002; Gruis et al. 2004). The expression pattern of *ASPA1* is in parallel with these VPEs and it may also process seed storage proteins, as studies showed that *ASPA1* accumulated in protein

storage vacuoles, and is able to cleave 2S seed storage protein napins *in vitro* (D'Hondt et al. 1993a; Mutlu et al. 1999; Otegui et al. 2006). The involvement of VPEs in seed germination is not documented, but the detection of ASPA in seed germination is reported (Pereira et al., 2008). It is possible that ASPAs are primary proteases in seed germination for degradation of seed storage proteins rather than the VPEs. This leads to another hypothesis that ASPAs regulate seed storage proteins procession during seed germination.

Among the three ASPAs in *Arabidopsis*, the most studied gene is *ASP3* due to its specific expression pattern in tissues that undergo programmed cell death (PCD), such as lateral root caps, tracheary elements in xylem, fading petals, tapetum in stamens and endosperm in developing seeds (Fendrych et al., 2014; Olvera-Carrillo et al., 2015). In plants, PCD can be categorized into two types: developmental PCD or environmental PCD (Daneva et al., 2016). Developmental PCD occurs in specific cell types such as tracheary elements in xylem, lateral root caps and tapetum in stamens, in order to facilitate normal growth and development. Developmental PCD also occurs in all types of aging cells in the end of plant senescence (Klimešová et al. 2015). Environmental PCD occurs in response to stresses such as irradiation or pathogens (Wu et al. 2014). Comparative studies showed that these two types are different in transcriptional signaling (Olvera-Carrillo et al. 2015). However, the executing components downstream of many PCD-related transcriptional factors are shared in different cell types (Olvera-Carrillo et al, 2015; Huysmans et al., 2018), such as *ASP3*,

BIFUNCTIONAL NUCLEASE1 (BFN1), *RIBONUCLEASE3 (RNS3)*, *CYSTEIN ENDOPEPTIDASE1 (CEP1)*, *DOMAIN OF UNKNOWN FUNCTION679 MEMBRANE PROTEIN4 (DMP4)* (Olvera-Carrillo et al., 2015; Ye et al., 2020). But the expression time is different for these genes. In *Arabidopsis* stigmas, the expression order is *CEP1* first, then *ASPA3*, and *BFN1* is the last (Gao et al., 2018). *CEP1* functions in the cytosol, and it is likely to participate in the signaling transduction. *BFN1* functions in the nuclei for the final degradation of DNA. This result indicates that *ASPA3* may function after the upstream signaling events and is involved in bulk proteolytic activity of cell components. In lateral root cap PCD, the final events include a decrease in cytoplasmic pH, plasma membrane permeabilization, vacuolar collapse, and final degradation cell materials (Fendrych et al. 2014). It is likely that *ASPA3* is one of the crucial proteases in these final steps, although the single mutant of *ASPA3* doesn't show a PCD-related phenotype in lateral root caps (Fendrych et al., 2014). It is likely that *ASPA1* and *ASPA2* have redundancy roles, and the third hypothesis can be drawn that *ASPAs* regulate programmed cell death in *Arabidopsis*.

A recent study showed another potential role of *ASPA1* in drought tolerance by overexpression in *Arabidopsis* (Sebastián et al., 2020). Overexpression lines of *ASPA1* had longer primary root length under drought conditions. The overexpressors also had reduced stomata index, reduced stomata density and a smaller stomatic aperture compared to wild type plants. Higher expression levels of genes related to ABA signaling and biosynthesis were also detected in *ASPA1* overexpression lines. *ASPA1*

promoter-GUS activity showed that *ASPA1* was induced by ABA in leaves. These results indicate multiple roles of ASPAs in stress responses.

ASPAs have been inferred as important proteases in plant growth and development, yet few studies have reported their activities *in vivo*. In this dissertation, three hypotheses mentioned above were tested, and the results showed that ASPAs function in seed storage protein processing *in vivo* during seed germination. ASPAs also regulate programmed cell death by membrane disturbance to increase membrane permeability in lateral root caps. These results indicate that ASPAs may regulate nitrogen storage and recycle in the plants.

Results

ASPAs function in seed development and germination

To explore the possible biological functions of ASPAs in *Arabidopsis*, T-DNA insertion mutants were used for phenotypic studies. There are three *ASPA* genes containing the saposin-like domain in *Arabidopsis*. Two alleles of *ASPA2* were obtained and characterized. The T-DNA insertions were verified by PCR. The *aspa2-1* (SALK097505) harbors a T-DNA insertion in the promoter and *aspa2-2* (SALK021601) has an insertion in the 5' untranslated region (Figure 2-01A). Quantitative real-time PCR revealed that both mutant lines are null alleles (Figure 2-01B). The single *aspa2* mutants showed delayed seed maturation reflected by the delayed seed size increase and delayed color change (Figure 2-02A). The delay corresponds to the heart stage of

development (Figure 2-02B). However, the phenotype was subtle, which suggests the possibility that other two *ASPAs* might have redundancy roles. T-DNA insertion mutants were also screened for *ASPA1* and *ASPA3*.

T-DNA insertion lines for *ASPA1* and *ASPA3* were obtained. The *aspa1-1* (SALK092586) has an insert in the 5' untranslated region, while *aspa1-2* (SALK041027) has an insertion in the second intron. The allele *aspa3-3* (SALK056711) has an insertion in the fifth intron, 5' of the PSI domain (Figure 2-02A). The *aspa1* alleles are both knock-down alleles with approximately 40% expression for wild type, and *aspa3-3* is a null allele (Figure 2-02B). The triple mutant *aspa1-2 aspa2-1 aspa3-3* (*aspa1-2/2-1/3-3* for short) was generated. A triple knockout could not be obtained, since *aspa1* alleles are knockdowns. Seeds of triple mutants were slightly larger and weighted almost twice as much compared to the wild type seeds (Figure 2-02C, D, E). This may result from delayed seed development allowing accumulation of more storage materials in the seeds. Considering the function of *ASPAs* in other species, *ASPA2* is likely involved in seed storage protein processing.

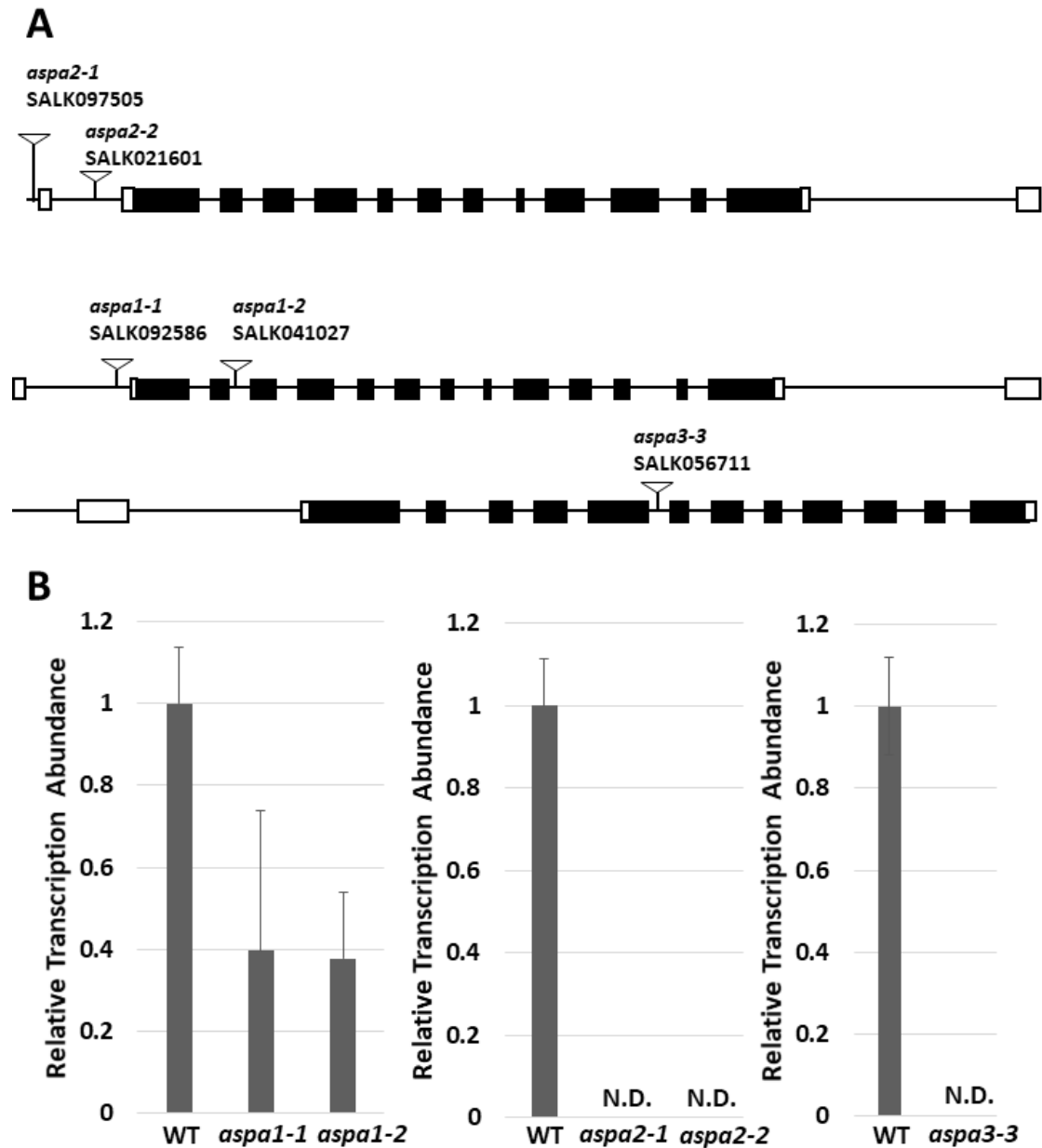


Figure 2-01. Characterization of ASPA T-DNA insertion mutants. (A) Schematic diagrams of the gene models and the T-DNA insertion sites in ASPA mutants. The inverted triangles represent the T-DNA insertion sites in the genomic DNA. Boxes represent the exons in the genomic DNA. Black boxes represent translated regions. White boxes represent untranslated regions. Intervening lines between black boxes represent the introns. (B) Quantitative real-time PCR of in T-DNA insertion mutants.

The inflorescence and opening flowers were harvested to extract RNA and quantitative real-time PCR was conducted to detect ASPA gene transcription level. ACTIN2 was chosen as an internal standard. Three biological replicates are represented in quantitative real-time PCR experiments. N.D.: none detected.

Seed germination was also affected in *aspa1-2 aspa2-1 aspa3-3* mutant. Seeds germinated more slowly than wild type seeds (Figure 2-03A), and this delay could not be rescued by gibberellin acid treatment (Figure 2-04A). This suggests that the delay may not result from transcriptional events. This also indicates that the major source of ASPAs during germination is synthesized during seed development, not newly synthesized after imbibition. Stratification resulted in alleviation in germination delay of mutant seeds (Figure 2-03A). This is probably because seed storage protein degradation was slower in the mutant due to lack of proteases and seed growth was slower as a result. To test this hypothesis, total proteins from imbibed seeds were extracted in a time course. By SDS-PAGE and Coomassie Blue staining, seed storage proteins were degraded faster in wild type and the protein levels decreased more slowly in mutant seeds (Figure 2-03B, C).

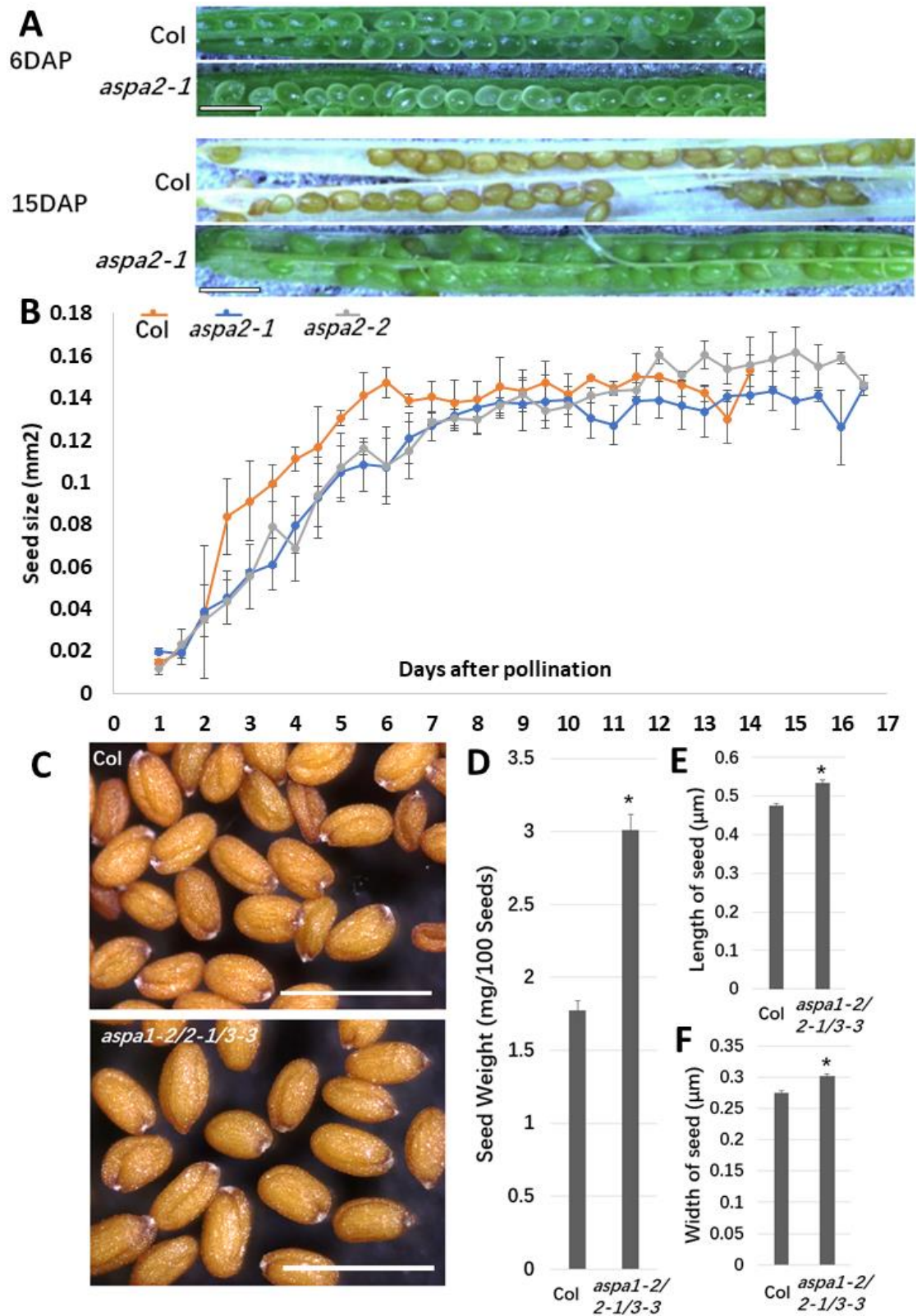


Figure 2-02. ASPAs regulate seed maturation in *Arabidopsis*. (A) Seed maturation was delayed in size increasing and color change in *aspa2* mutant seeds. Representative images show the relative seed size and color of wild type and mutant seeds for 6 DAP

(top) and 15 DAP (bottom). DAP: days after pollination. Bar=1mm. (B) Rate of seed size increases over time in wild type and *aspa2* mutant. Five individual plants were selected and for each plant at least ten developing seeds were measured at each timepoint for statistical analysis. (C) Representative images of wild type and *aspa1-2 aspa2-1 aspa3-3* seeds. Bar=1mm. Seeds were freshly harvested and put in drying oven at room temperature for at least three days, within a month. (D) Weight of wild type and *aspa* mutant fully mature seeds. Three biological replicates were represented and for each replicate, 200-300 seeds were weighed. Asterisks indicate statistical significance $p < 0.05$ for Student' *t*-test. (E) Seed length and (F) seed width of wild type and *aspa* mutants. N=50 seeds were selected for three replicate each. Asterisks indicate statistical significance $p < 0.05$ for Student' *t*-test.

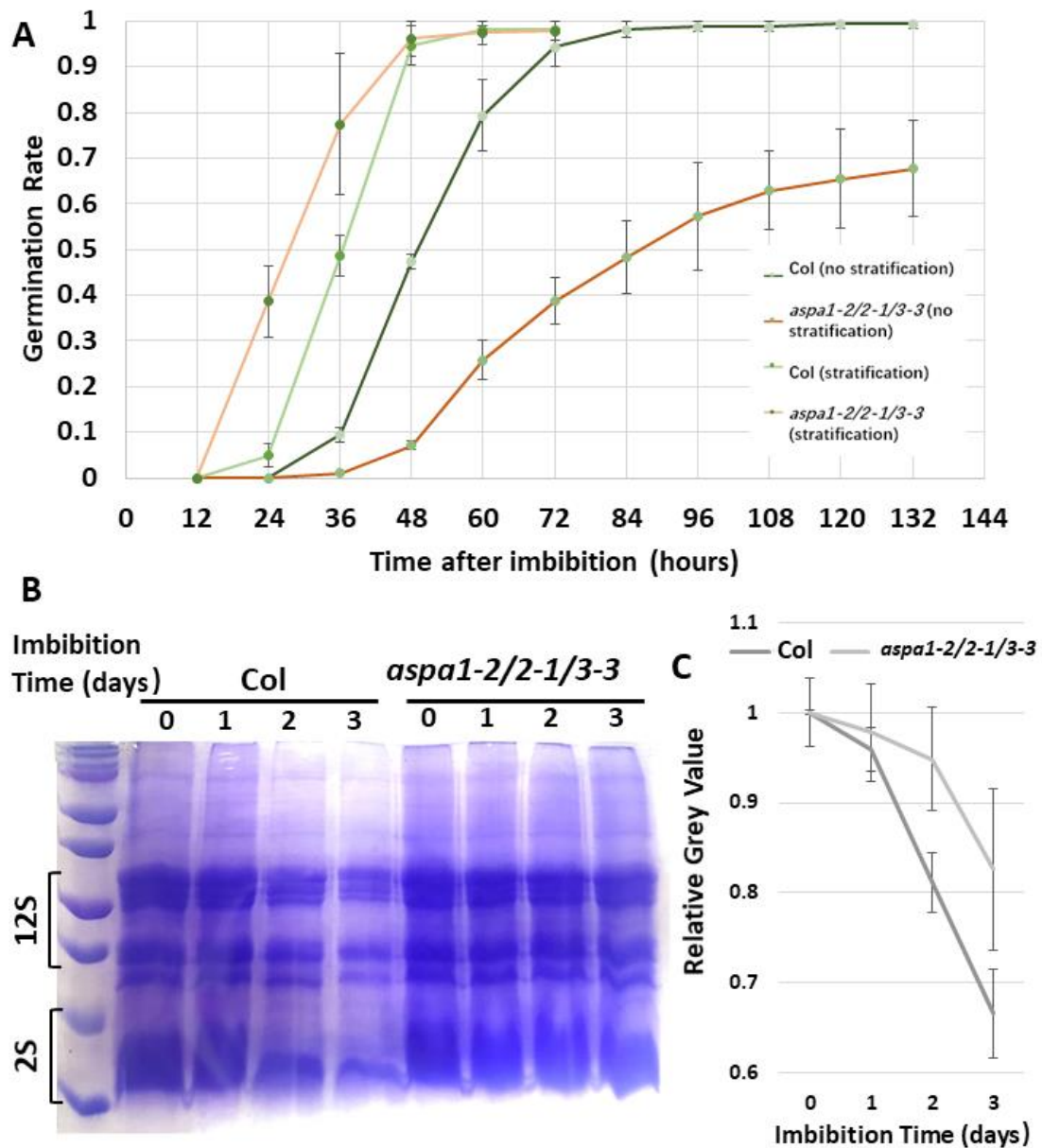


Figure 2-03. Seed germination and seed storage protein degradation in *aspa* mutant.

(A) Germination rates of Col-0 and *aspa1-2/2-1/3-3* seeds germination rate with and without stratification. The delayed germination in *aspa* mutant seeds could be rescued by stratification for 2 days. For germination experiments, three biological replicates and N=150 fresh mature seeds were selected for each line for each replicate. (B) Coomassie blue staining of seed storage proteins from imbibed *aspa1-2/2-1/3-3* and wild type seeds. Total proteins of approximately 20 seeds were loaded for each lane. (C) Relative grey value of seed storage proteins from imbibed *aspa1-2/2-1/3-3* and wild type seeds. Total proteins of approximately 20 seeds were loaded for each lane.

(C) Relative intensity of staining in B was measured with ImageJ. Time 0 was chosen as baseline and set to 1.

In addition to new protein synthesis, cell expansion is also an important aspect in increasing cell volumes during seed germination. A major factor contributing to cell expansion is the central vacuole fusion and expansion. During seed germination, the small protein storage vacuoles fuse with each other and form the large central vacuole. By absorbing lots of water during imbibition, central vacuoles increase in volume, the embryo rapidly expands, the radicle breaks the seed coat and grows into the soil. To test whether storage protein vacuolar fusion is also delayed in the mutants, the embryos were dissected from the imbibed seeds over a time course. The embryo cotyledons were visualized by autofluorescence with confocal microscopy (Figure 2-05A). The protein storage vacuoles fusion was slower in the mutant cells compared to the wild type. This indicates that ASPAs may also regulate membrane disturbance for vacuolar fusion during seed germination. Total protein was also extracted from imbibed seeds and analyses via SDS-PAGE followed by Coomassie blue staining. The protein gel analysis showed that the protein degradation was slower in the mutant seeds compared to the wild type (Figure 2-03B, C).

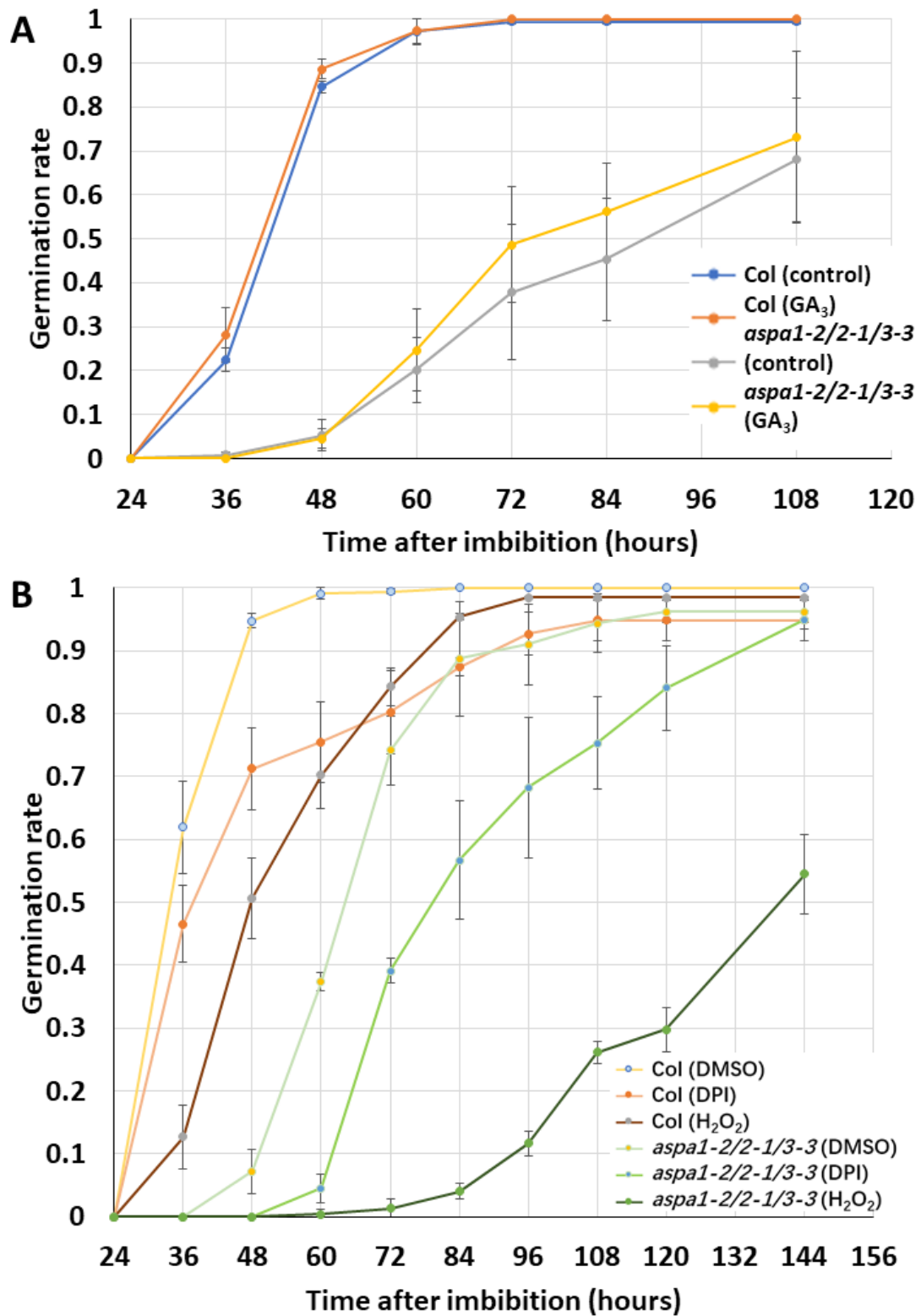


Figure 2-04. Germination rates in Col-0 and *aspa* triple mutant seeds with and without gibberellin acid 3, hydrogen peroxide and diphenylene iodonium treatments. (A) Germination rates of Col-0 and *aspa* mutant seeds with and without gibberellin acid 3

(GA₃) treatment. Seeds were sown on 1/4MS media supplemented with 1μM GA₃. Three biological replicates with N=150 seeds for each line for each replicate. (B) Germination rates of Col-0 and *aspa* triple mutant seeds with and without hydrogen peroxide (H₂O₂) or diphenylene iodonium treatment. Seeds were sown on 1/4MS media supplemented with 10mM H₂O₂, 10μM diphenylene iodonium (DPI) and 0.2% DMSO (solvent) for control.

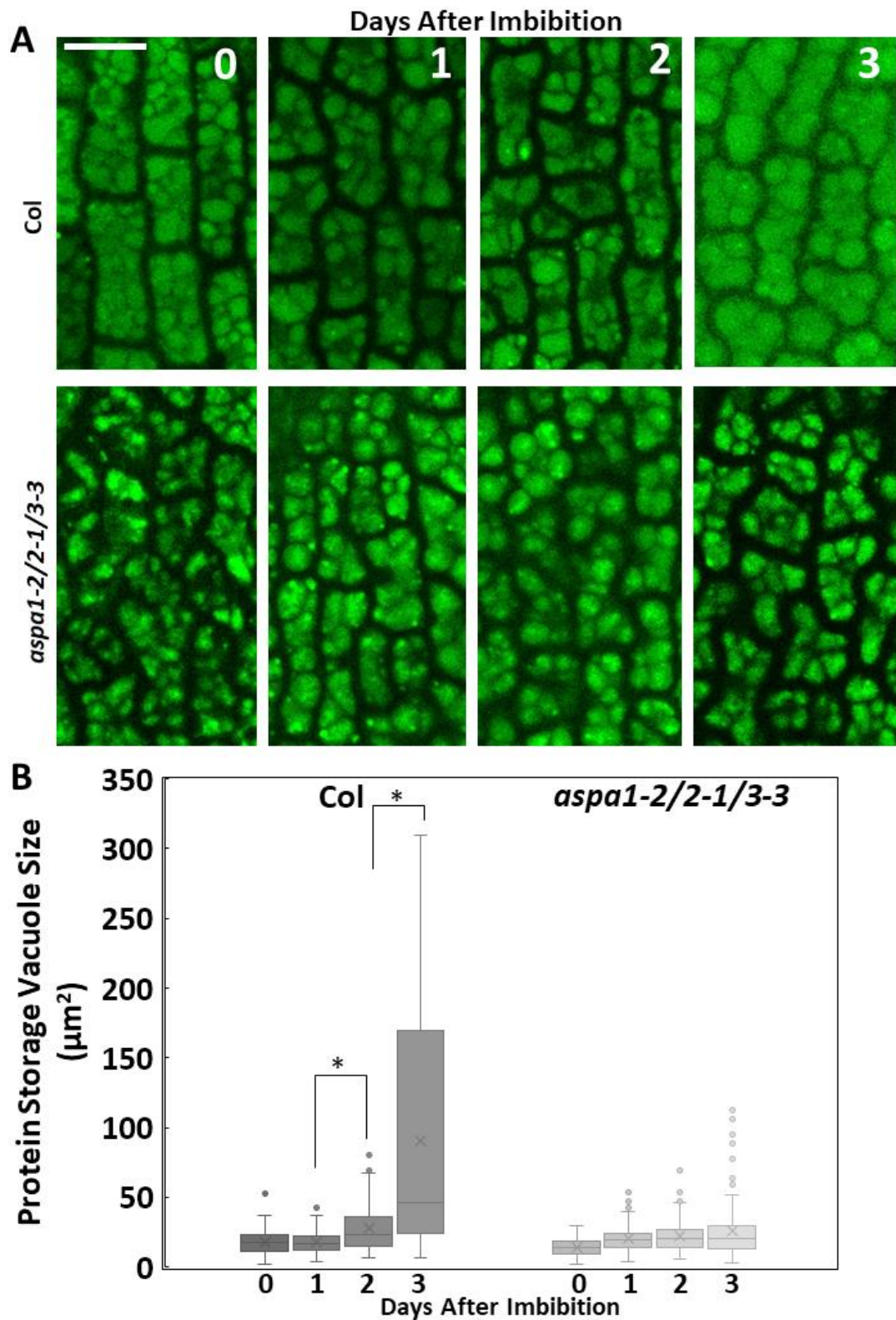


Figure 2-05. Protein storage vacuole (PSV) fusion during germination in Col-0 and *aspa* triple mutant seeds. (A) Morphology of protein storage vacuoles in imbibed wild type and *aspa1-2/2-1/3-3* mutant seeds. Autofluorescence in cotyledons was imaged at

488 nm excitation. Bar=20µm. (B) Protein storage vacuole sizes in Col-0 and aspa triple mutant seeds over time. N=14-20 for each line for each replicate at each time point. Asterisks indicate statistical significance, $p < 0.05$, ANOVA followed by Tukey post-hoc test with six replicates.

These results suggest that ASPAs are directly involved in seed storage proteins processing, and function downstream of signaling during seed maturation and germination. They take part in metabolism rather than signaling. If this is the case, the mutant seeds should be more sensitive to environment stress that affects metabolism. To test this hypothesis, seeds were treated with hydrogen peroxide to mimic reactive oxygen species stress in overly active metabolic state in the cell, and with NADPH synthase inhibitor diphenyleneiodonium to reduce hydrogen peroxide production, to mimic inhibited metabolic state in the cell. The mutant seeds were more sensitive these treatment as the germination rate was even slower (Figure 2-04B). In summary, ASPAs are involved in seed maturation and seed germination by processing seed storage proteins. The sensitivity to environmental stress may be disadvantage trait in evolution selection.

Expression pattern of *ASPA2*

As previously published, *ASPA1* mRNA is detected in all tissues and more abundant in leaves during daytime. *ASPA3* is primarily in flowers and *ASPA2* is primarily

in seeds. (Chen et al., 2002; Sebastián et al., 2020). The *ASPA3* promoter-reporter constructs showed signals in almost all tissues that undergo programmed cell death (PCD), such as lateral root caps, tracheary elements in proxylem, fading petals, tapetum in stamens and endosperm in developing seeds (Fendrych et al., 2014; Olvera-Carrillo et al., 2015). In terms of *ASPA2*, no reports have been shown about its expression in other tissues except seeds. To test the expression pattern of *ASPA2*, 2kb promoter was cloned and incorporated into an expression construct with the reporter HISTONE 2A 10 (H2A) fused to a YFP tag. The expression appeared in the suspensor at the globule stage (Figure 2-06B). In developing embryos, *ASPA2* was expressed beginning at heart stage (Figure 2-06C) throughout seed maturation (Figure 2-06D, E). It was also expressed in integuments and endosperms (Figure 2-06D).

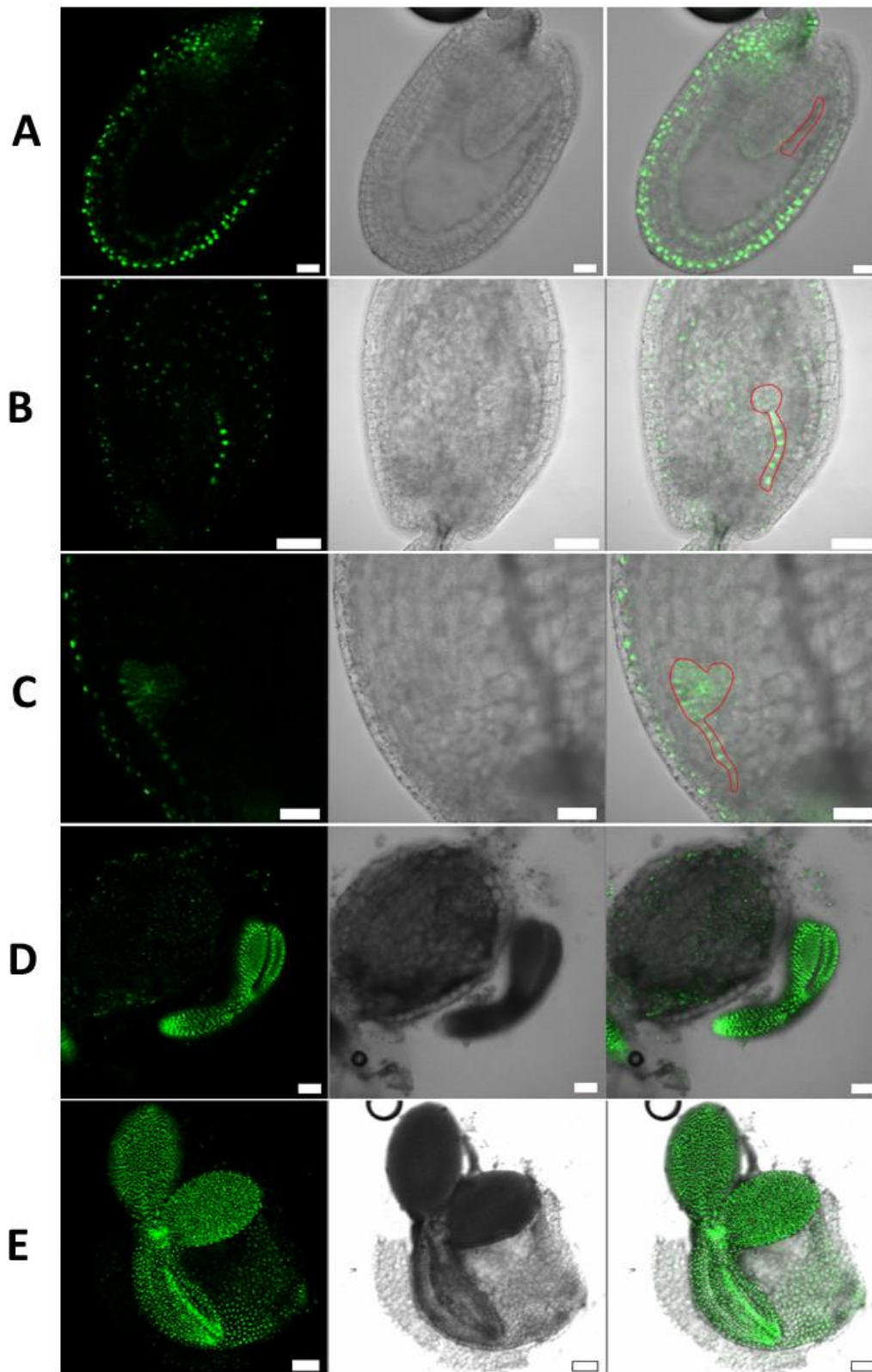


Figure 2-06. *ASPA2* expression during seed development. Embryo is depicted with red outlines. *ASPA2* promoter:: H2A-YFP reporter lines in Col-0 background were imaged using confocal laser scanning microscopy. (A) Early globular stage. Bar=20 μ m. (B)

Globular stage. Bar=50µm. (C) Early heat stage. Bar=50µm. (D) Torpedo stage. Bar=50µm. (E) Mature stage. Bar=100µm. Embryo is outlined in red.

In seedling and vegetative tissues, the expression was detected in almost all tissues: roots, hypocotyls, cotyledons, true leaves (Figure 2-07A to D). In reproductive tissues, signals were detected in stems, sepals, stamens including filaments and the anther epidermis (but not pollen), carpels, stigma, transmission tissues and ovules (Figure 2-07 E to J). In general, *ASPA2* was ubiquitously expressed in *Arabidopsis* plants, similar to *ASPA1*. This suggests that *ASPA2* is likely to be redundant with *ASPA1*. It also indicates that ASPAs have other functions besides seed maturation and germination. The functions in other tissues remains unclear and need to be explored.

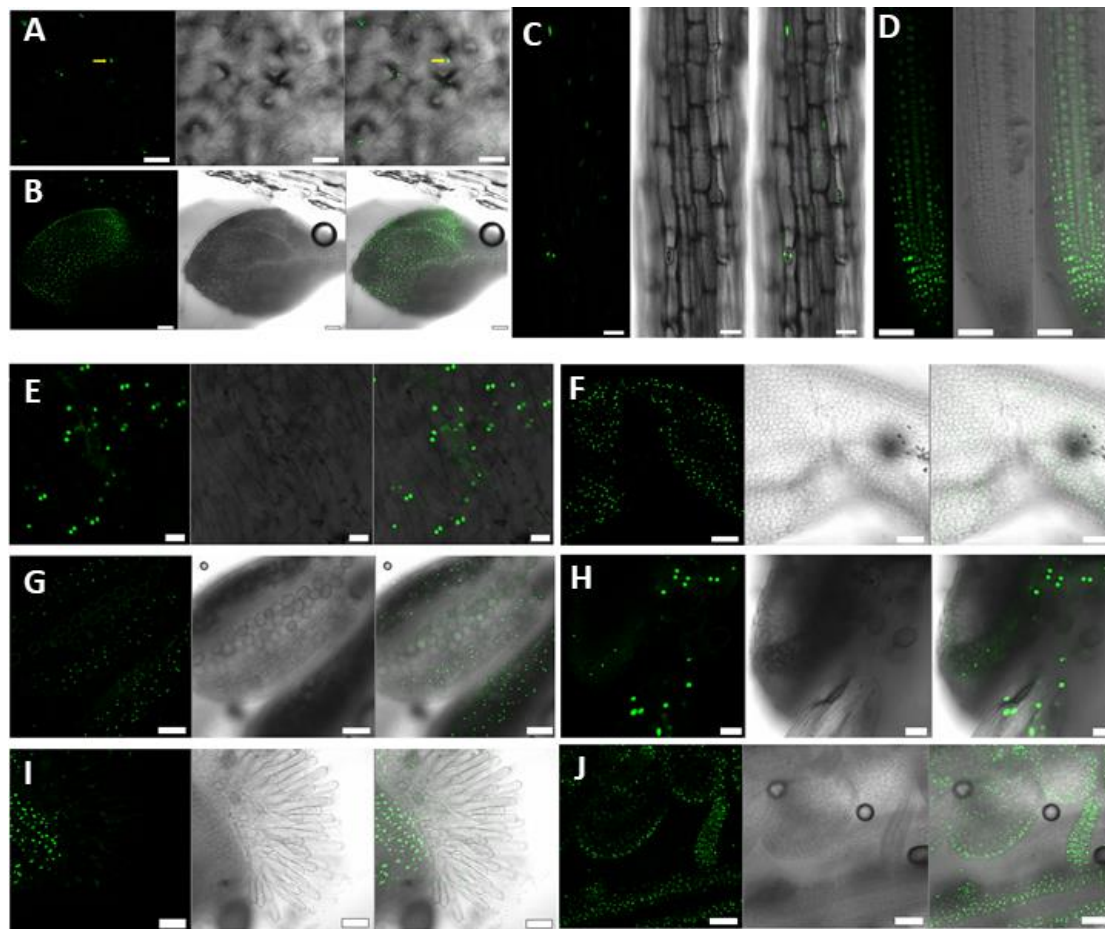


Figure 2-07. *ASPA2* expression in vegetative and reproductive tissues. *ASPA2* promoter:: H2A-YFP reporter lines in Col-0 background were imaged using confocal laser scanning microscopy. (A) Cotyledon. The yellow arrow points to nucleus in the epidermis. Bar=20 μ m. (B) First pair of true leaves. Bar=20 μ m. (C) Hypocotyl. Bar=20 μ m. (D) Root. Bar=20 μ m. (E) Sepal. Bar=20 μ m. (F) Petal. Bar=50 μ m. (G) Anther. Bar=50 μ m. (H) filament. Bar=20 μ m. (I) Stigma and carpel. Bar=50 μ m. (J) Transmission tissue. Bar=50 μ m. For all images, left: YFP; middle: TML, transmitted light; right: merge.

Subcellular localization and trafficking of ASPA2

To further explore the functions of ASPA2 in other tissues, overexpression plants

were generated. Coding sequence (CDS) of *ASPA1*, *ASPA2* and *ASPA3* were cloned and inserted into plant expression vector driven by 35S promote fused with either the CFP-HA tag or the RFP tag on the C-terminus. The constructs were transformed into wild type background. Full length proteins were detected by Western blotting (Figure 2-09A).

The subcellular localization and trafficking of *ASPA2* was analyzed. *ASPA2* was trafficked to the vacuoles (Figure 2-08B). *ASPA1* and *ASPA3* were also trafficked to vacuoles (Figure S07). With brefeldin A (BFA) treatment, a fungal inhibitor which blocks trafficking between endoplasmic reticulum (ER) and Golgi complex, *ASPA2* colocalized with *trans*-Golgi body network (TGN) marker SYP61 (Figure 2-08A) and TGN/early endosome (EE) marker RabA3 (Figure 2-08B). With concanamycin A (conc A) treatment, which inhibits the vacuolar type H-ATPase and further inhibits fusion with vacuoles, *ASPA2* colocalized with the multivesicular body (MVB)/prevacuolar compartment (PVC) marker RabF1/ARA6 (Figure 2-08C). These results showed that *ASPA2* is first synthesized on ER, transported to TGN and then trafficking to MVB/PVC, finally to the vacuoles. The route passing through the TGN and fuse with EE compartments suggests that proteins on plasma membrane may also contact with *ASPA2*.

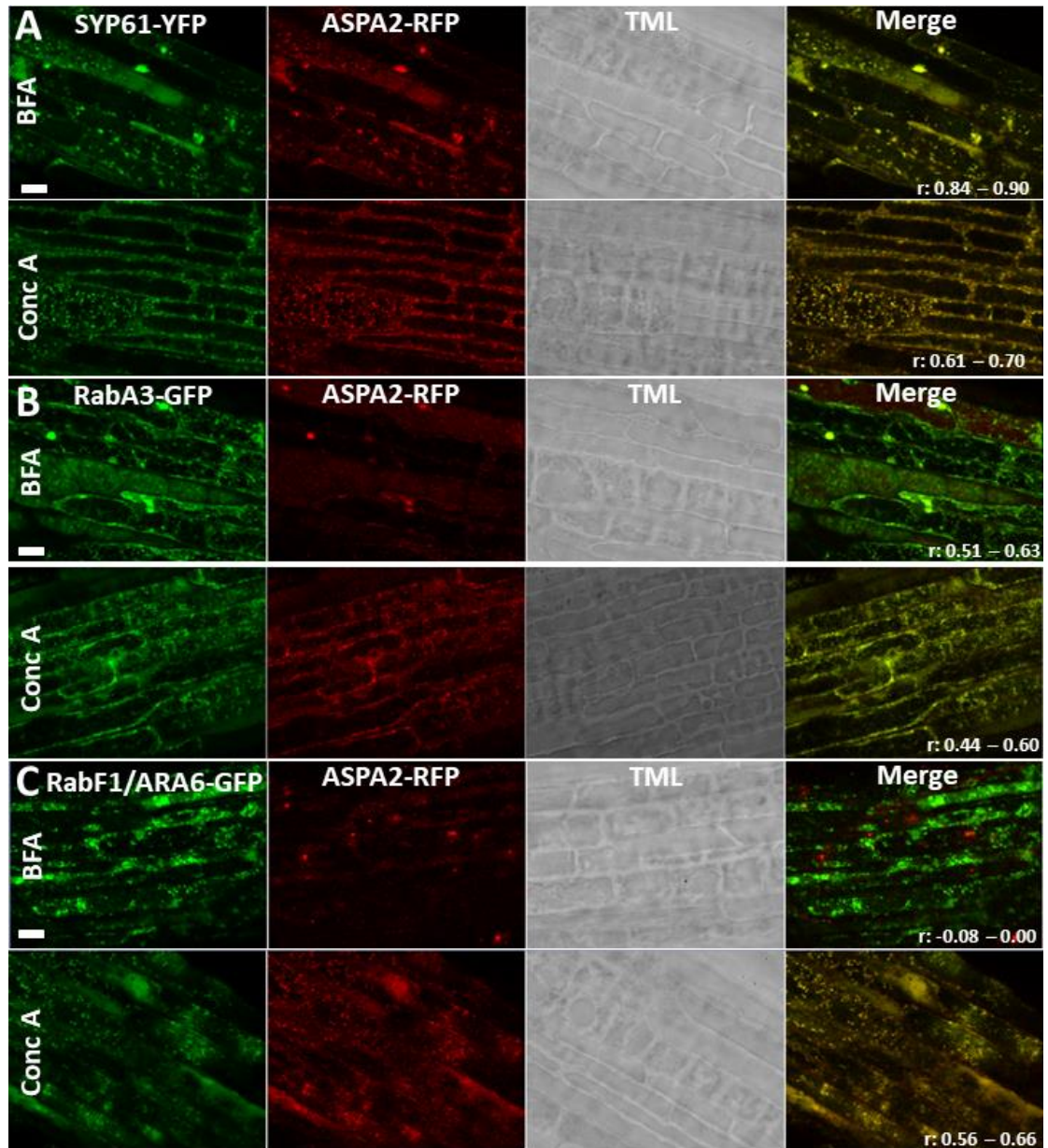


Figure 2-08. Intracellular trafficking pathway of ASPA2 in *Arabidopsis* roots. Colocalization of 35S:: ASPA2-RFP in Col-0 background with TGN marker SYP61-YFP (A), TGN/EE marker RabA3-GFP (B) and MVB marker RabF1/ARA6-GFP (C). Images were taken after one hour incubation with 10 μ M brefeldin A (BFA) or 100nM concanamycin A (Conc A). ASPA-RFP did not colocalize with RabF1/ARA6-GFP with BFA treatment. Bar=10 μ m. Pearson's correlation (r) range is listed in each panel with three seedlings.

The saposin-like domain, or plant specific insert (PSI) has been suggested as vacuolar trafficking signal (Kervinen et al., 1999; Terauchi et al., 2006). The PSI deletion version of ASPA2 was generated in this dissertation, which was 35S promoter::ASPA2-321AA-CFP-HA construct. This deleted ASPA2 contains 1st to 321st amino acid residues. Result showed that there were signals failing to traffic to vacuoles (Figure 2-09B). This indicates the PSI is required in vacuolar targeting.

ASPAs are processed to produce mature enzymes. To test whether self-catalytic activities are required for vacuolar trafficking, the first conserved aspartic site was mutated to the alanine in this dissertation, which was the 35S promoter::ASPA2-D107A-CFP-HA construct. This mutation abolishes protease activity, and no self-proteolytic activity occurs. Result showed that ASPA2-D107A-CFP were trafficked to vacuoles. This indicates that proteolytic activity is not required for vacuolar trafficking. This catalytic inactive protease version keeps the intact PSI, and it could be regarded as a PSI overexpression in plants for further analysis.

Studies show that conformation change is important for saposin-like proteins interacting with lipids. A novel six-amino acid-motif [N/Q]-[N/Q]-[A/L/I/V]-[K/R]-[N/Q] in helix H3 of the saposin-like domain from the potato aspartic protease StAP appeared to be responsible for interaction with membrane lipids, as a point mutation blocked the conformational change and abolished the membrane fusion ability *in vitro* (Bryksa et al., 2017). Conformation change has been reported for saposin-like proteins interacting with membrane lipids for human saposin C and D (Rossmann et al., 2008).

It could be speculated that this motif affects PSI function in vacuolar targeting. If conformation change is required for vacuolar targeting, a point mutation in this motif may block aspartic protease trafficking to vacuoles. To test whether the abolishment of this motif in saposin-like domain of ASPA2 affects vacuolar targeting, the point mutation was generated in this motif (R402Q) in the ASPA2-D107A context. Which was 35S promoter:: ASPA2 D107A R402Q-CFP-HA in this dissertation. Results showed that both versions showed the vacuolar subcellular localization (Figure 2-09B). This result indicated that this motif is not related to vacuolar targeting of ASPA2. This also suggests that vacuolar targeting role of PSI doesn't require conformation change in the cell.

The modification of PSI was also investigated. There's only one potential glycosylation site (N404) in ASPA2, which is just after the six-amino acid motif. Glycosylation may also affect the interaction between the motif in PSI and the membrane lipids. To test whether this glycosylation affects ASPA2 vacuolar targeting, the point mutation version 35S promoter:: ASPA2 D107A N404A-CFP-HA was generated and transformed into *Arabidopsis*. Results showed that ASPA2 D107A N404A-CFP targeted in the vacuoles. This suggests that glycosylation doesn't affect vacuolar targeting either (Figure S04).

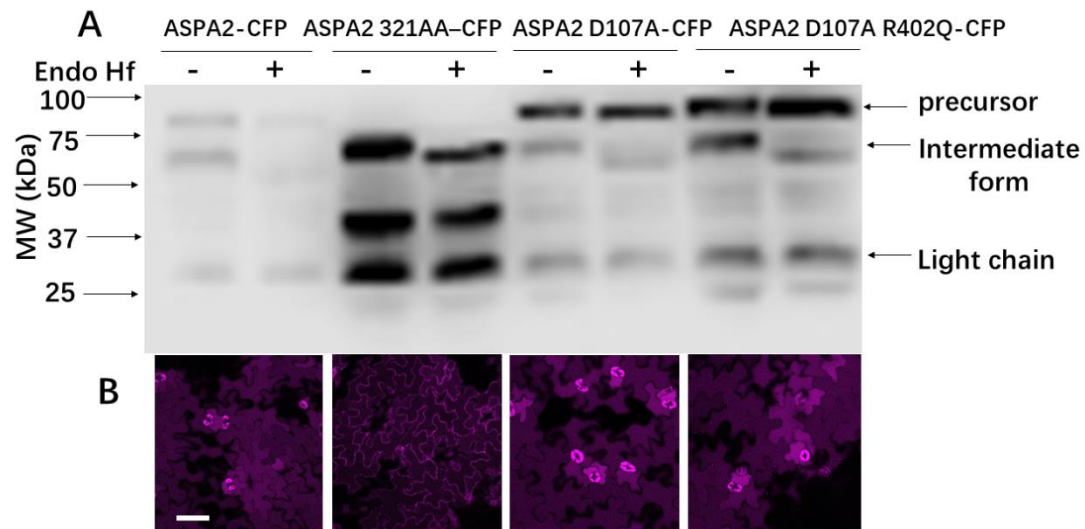


Figure 2-09. Glycosylation and vacuolar trafficking of ASPA2-CFP. (A) Western blot analysis of total proteins from 30mg mature leaves for each lane. Samples were treated with/without (+/-) Endo Hf to evaluate glycosylation of ASPA2-CFP: native ASPA2, deletion of PSI and C-terminus version (ASPA2 321AA), single point mutation version in conserved aspartyl site (ASPA2 D107A) and double point mutations in conserved aspartyl site and mutation in lipid binding motif (ASPA2 D107A R402Q). (B) Confocal laser scanning images of the subcellular localization of these ASPA2 mutations in cotyledons of 14-day-old seedlings. Bar=100µm. 10-20 seedlings were tested.

ASPAs are involved in root architecture regulation

To further elucidate the biological functions of ASPAs in other tissues, the triple mutants and overexpression lines were grown for phenotypic analysis. The primary root of seedlings was slightly shorter in 35S ::ASPA2-RFP overexpression lines than wild

type (Figure 2-10A, B). The overexpression level of ASPAs were not verified in this dissertation, but the full-length protein and fluorescent tag was verified (Figure 2-09). While there's no significant difference in primary root growth length between triple mutant and wild type, more lateral roots formed in the mutants (Figure2-13C).

Since ASPAs are likely to involved in seed storage protein processing, this suggests that ASPAs are important in nitrogen metabolism in *Arabidopsis*. To test whether ASPAs are also involved in nitrogen metabolism in vegetative tissues, seedlings were transferred to low nitrogen media. The primary root length in the mutants was longer than wild type, and the mutant roots were insensitive to the low nitrogen treatment (Figure 2-10B). The overexpression lines did not show significant differences from the wild type (Figure 2-10B). This suggests that ASPAs affect root architecture in *Arabidopsis* with respect to integration of nutritional signals.

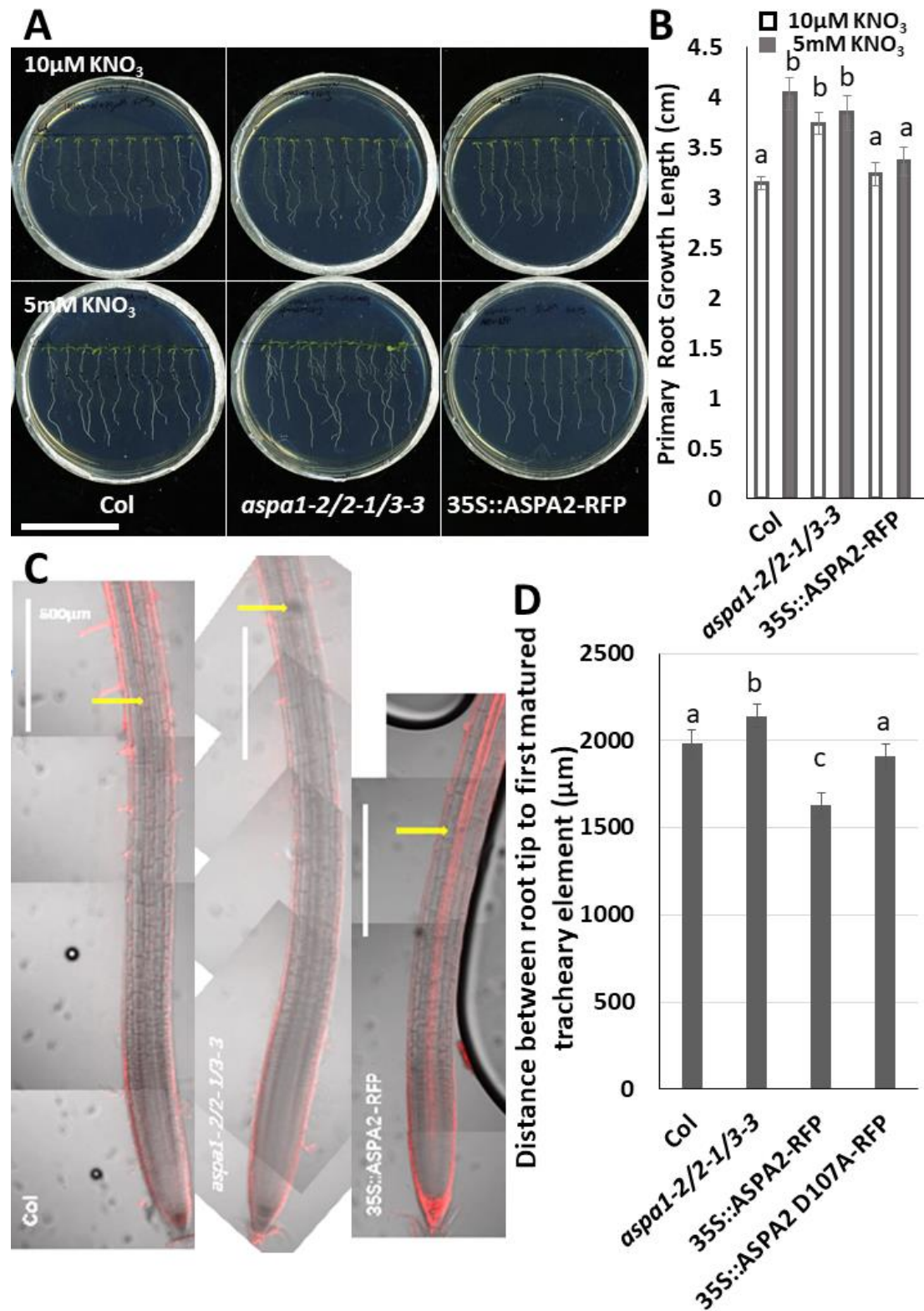


Figure 2-10. Root architecture in Col-0 and *ASP2* mutants. (A)-(B) Primary root growth in response to low nitrogen in wild type, *aspa1-2/2-1/3-3* mutant and 35S:: *ASP2*-RFP seedlings. Bar=5cm. Seedlings were grown on regular 1/4MS media for 4 days and then

transferred to media with low nitrogen (10 μ M KNO₃) or sufficient nitrogen (5mM KNO₃) for additional 4 days. Black dots mark the position of 4DAG seedling root tip at the time of transfer. (B) Statistics of primary root growth length in Col-0, *aspa1-2/2-1/3-3* mutant and 35S::ASPA2-RFP seedlings. Different letters indicate statistical significance, $p < 0.05$ ANOVA followed by Tukey post-hoc test. (C)-(D) Treachery elements maturation in Col-0, *aspa1-2/2-1/3-3*, 35S::ASPA2-RFP seedlings. (C) Representative image of 7 DAG seedlings stained with 4 μ M propidium iodide. Yellow arrows indicate the first matured treachery elements. Bar=500 μ m. (D) Distance between root tip and the first matured treachery elements. N=10-15 seedlings. Different letters indicate statistical significance, $p < 0.05$ ANOVA followed by Tukey post-hoc test.

Transcriptional regulation of *ASPA2*

To test whether *ASPA1* and *ASPA2* are transcriptionally regulated by environmental signals, the promoters of *ASPA1* and *ASPA2* were cloned and incorporated in expression constructs with reporter HISTONE 2A 10 (H2A) fused with either mCherry or YFP. The reporter lines were treated with low nitrogen, abscisic acid (ABA) and sodium chloride (NaCl). There are ABA responsive *cis* elements in *ASPA2* promoter, while none is found in *ASPA1* promoter. *ASPA1* expression was not changed with these treatments (Figure 2-11A, B). *ASPA2* expression was slightly higher (approximately 1.5 times higher) with low nitrogen treatment, highly upregulated (approximately 3 times higher) by ABA, and downregulated (approximately 60% lower)

by NaCl. This suggests that *ASPA2* is responsive to different environmental signals, while *ASPA1* functions like a housekeeping gene.

However, the primary root length was affected to the same extent with either ABA or NaCl treatment compared to the wild type, except that *ASPA1* overexpression plants showed slightly shorter roots with ABA treatment. This might be the artificial effects of *ASPA1* overexpression in the plants, or the genes were not highly overexpressed in the plants.

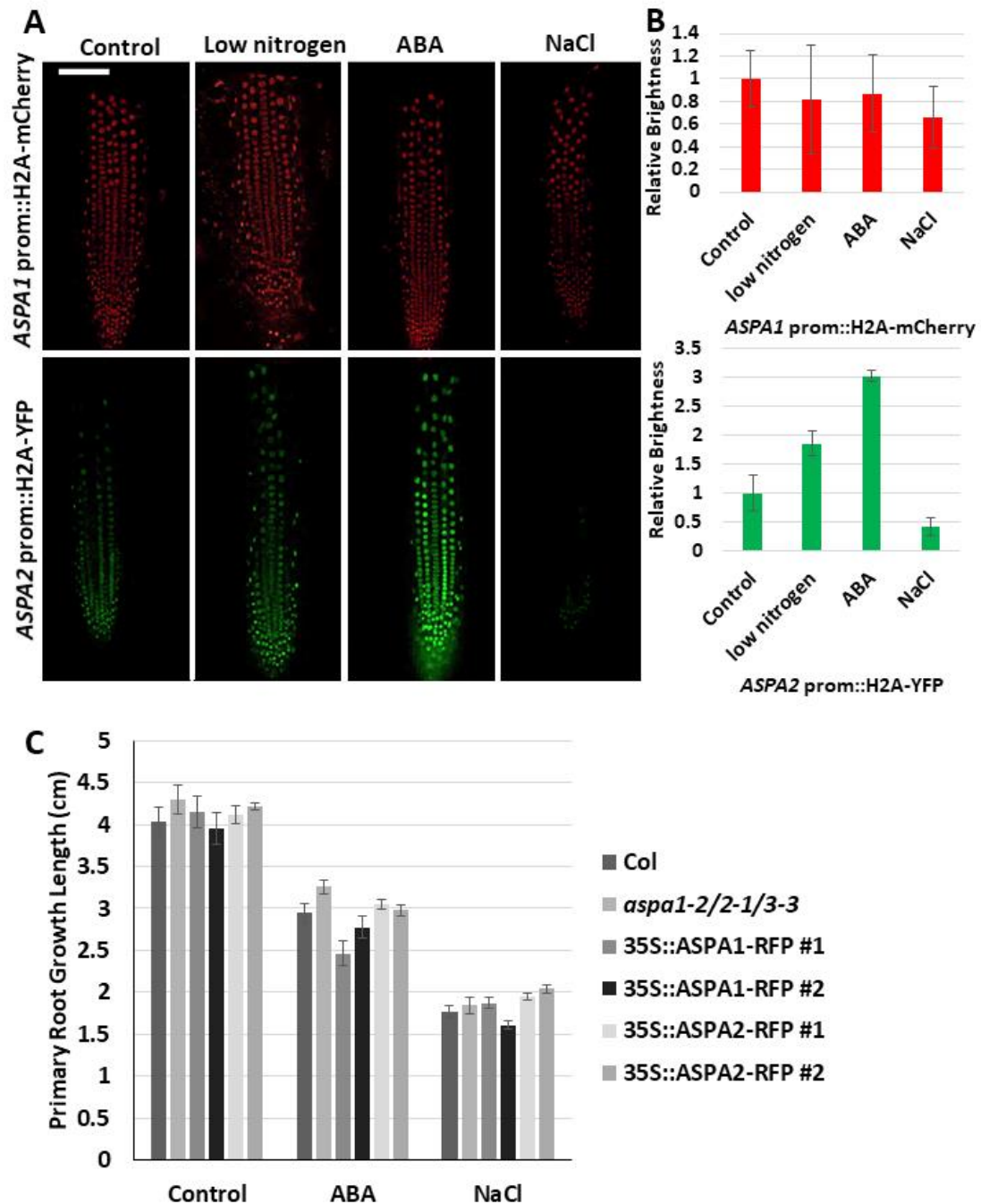


Figure 2-11. Transcriptional regulation of *ASPA1* and *ASPA2* and root growth in responses to ABA and NaCl treatments. (A) Confocal laser microscopy images of *ASPA1*::H2A-mCherry (Top) and *ASPA2*::H2A-YFP (bottom) reporter lines in responses to low nitrogen (10 μ M KNO₃), ABA (2 μ M) and NaCl (75mM) treatment. Seedlings were growing on 1/4MS media for 5 days and then transferred to new plates containing the

corresponding chemicals for another 2 days. 0.2% ethanol (solvent) as the control. Bar=100µm. N=50 nuclei from 5 seedlings in each line, each treatment. (B) Statistics of the fluorescent intensity based on (A). (C) Primary root growth length of Col-0, *ASPA1* overexpression lines and *ASPA2* overexpression lines with ABA (2µM) and NaCl (75mM) treatments. Seedlings were growing on 1/4MS media for 4 days and then transferred to new plates containing the corresponding chemicals for another 4 days. 0.2% ethanol (solvent) as the control. N=10 seedlings with three replicates.

To further examine how root architecture was affected, the position of the first mature tracheary element was measured, since disturbance of xylem maturation affects root growth in plants. 7DAG seedlings were treated with propidium iodide (PI) to visualize the spiral pattern of tracheary elements, and the distance between the first tracheary element and root tip was measured. This distance was slightly longer in triple mutant roots, and slightly shorter in *ASPA2* overexpression plant roots (Figure 2-10C, D). These results suggest that xylem maturation was slightly slower in triple mutant and slightly faster in overexpression plants. When the first conservative aspartic site in the protease was mutated (D107A), the distance was not affected. This indicates that the proteolytic activity is necessary for xylem maturation. Since *ASPA3* has been believed to take part in programmed cell death (PCD) in *Arabidopsis*, although *aspa3-3* single mutant doesn't show PCD related phenotype in lateral root cap cells in the published work (Fendrych et al., 2014), it is likely that all three ASPAs

are involved in regulation of PCD of tracheary elements and thus affect root morphology.

Autophagy pathway is activated under nutrient deficient condition and autophagy is also involved in programmed cell death (PCD). It is possible that ASPAs are associated with autophagy pathway in response to low nitrogen supply and PCD in TE. To test whether ASPA2 is involved in autophagy pathway, colocalization between autophagy marker ATG8a and ASPA2-RFP was imaged. Colocalization was not found with concanamycin A treatment (Figure S01A). This suggests that ASPA2 is not associated with autophagy pathway. A dual functional endosomal sorting complex required for transport (ESCRT) machinery associated protein FREE1 is reported to regulate both autophagy pathway and MVB formation (Gao et al., 2015). Colocalization results showed partial colocalization between GFP-FREE1 and ASPA2-RFP (Figure S01B). This indicates that ASPA2 colocalized with FREE1 in MVB, not in the autophagy compartments. These results also indicate that the PCD type in which ASPA2 is associated with is apoptosis type rather than autophagy type.

ASPAs are involved in programmed cell death

To test whether ASPAs might be involved in programmed cell death, the lateral root cap was chosen as a model system to study because it is easier to observe and the whole process can be monitored. The dead cells were identified by propidium iodide staining of the nuclei. Two measurements will indicate the initiation/onset of

PCD and the rate of PCD. The distance between root tip and all stained nuclei indicates the onset of PCD. The distance between root tip and all stained nuclei was measured (Figure 2-12 A). There was no difference in this distance distribution between the triple mutant and wild type. This suggests that the onset of PCD was not different between the mutant and wild type.

Then the distance between the distal stained nucleus and root tip was measured to determine the rate of PCD. If the cell collapses and peels off, there would be no signal. Since cell division continues at root tip at the same rate, if the cell collapse is slower and delayed in peeling off, then this distance would be longer. The results showed that this distance was longer in triple mutant (Figure 2-12B, C). This suggests that PCD execution process was slower in triple mutant. By monitoring the appearing and disappearing PI signals, the triple mutants showed a longer sustained PI signal over the time period, which suggests that cell death was slower in the triple mutant (Figure 2-12D). The PCD onset in the mutant and wild type were not different but the rate of cell death was different. This indicates that ASPAs may function in degrading cell components rather than signaling transduction during PCD.

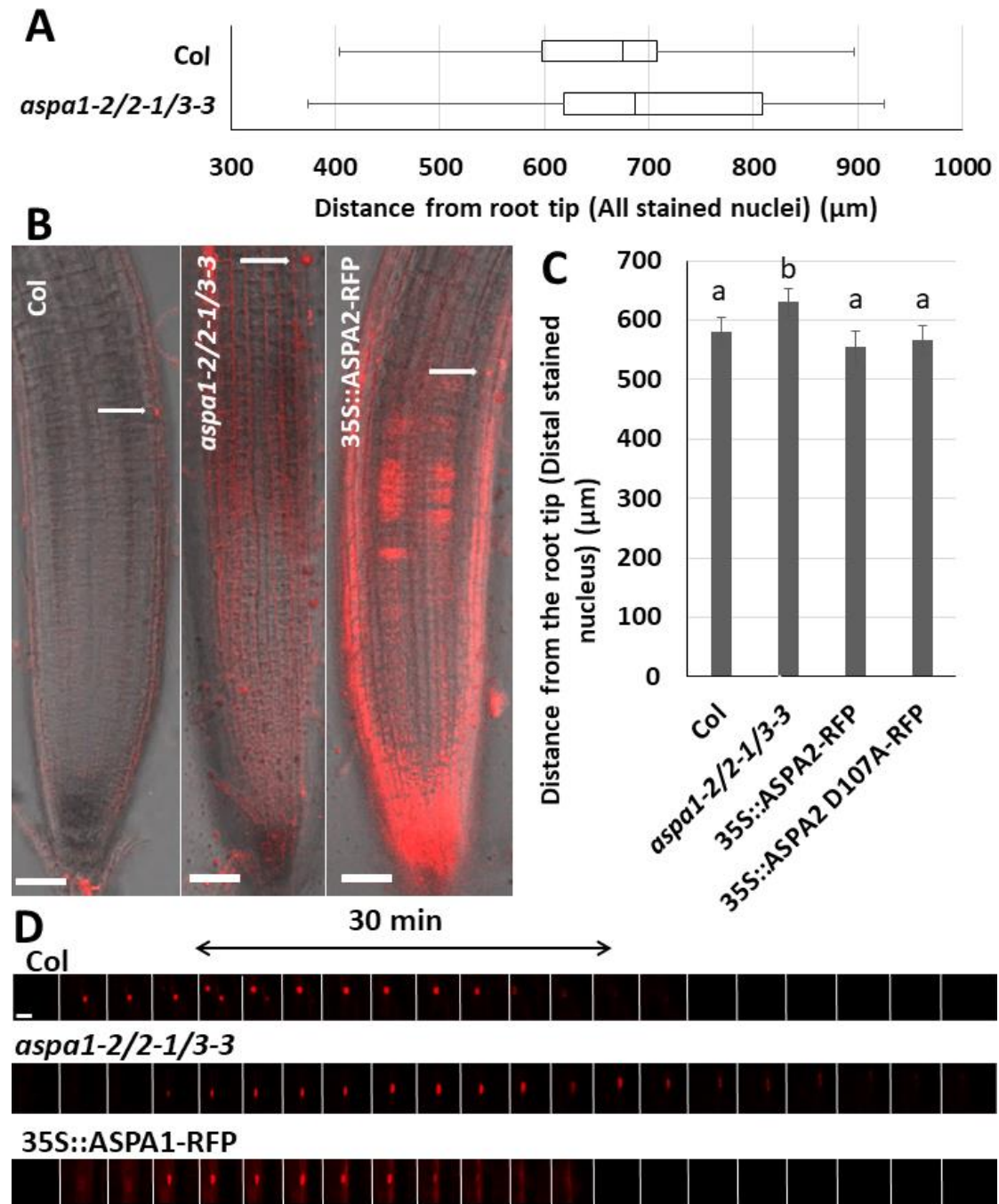


Figure 2-12. Propidium iodide (PI) staining in lateral root caps of Col-0 and *ASPA* mutants. (A) Distance between the root tip and all stained nuclei in Col-0 and *aspa1-2/2-1/3-3*. 6 days after germination (DAG) Col-0 and *aspa1-2/2-1/3-3* seedlings were stained with 4μM PI. N=20-40 nuclei from 5-8 seedlings per line were imaged. $P>0.05$ by Student' *t*-test. (B) Representative images of 7 DAG seedlings stained with 4μM

propidium iodide. White arrows point to the distal cell stained with 4 μ M PI. Bar=20 μ m.

(C) The distance between root tip and the distal nucleus stained with PI. 8-12 seedlings per line were imaged and measured. Different letters indicate statistical significance, $p < 0.05$ by Tukey post-hoc test with ANOVA. (D) PI stained nucleus in Col-0, *aspa1-2/2-1/3-3* and 35S::ASP1-RFP lateral root caps over time. Bar=20 μ m. 6DAG seedlings were stained with 4 μ M PI and imaged every five minutes.

To further investigate the possible roles of ASPAs in PCD processes, fluorescein diacetate (FDA) was used to stain living cells. The dye emits green fluorescence in the cytosol under neutral pH, and the intensity drops dramatically with decreasing pH. During PCD there is a pH drop in the cytosol, and this is considered as one of the first events of PCD. Then cell membrane permeability increases, and DNA is fragmented, cell components are compartmented afterwards (Fendrych et al., 2014). Therefore PCD could be indicated by the disappearance of FDA signals. Then with the increasing permeability of cell membranes, PI enters cell and stains the nucleus. The time period between the pH decrease and PI staining indicates the rate of disruption of the cell membrane system and increasing cell membrane permeability. This time period was monitored by FDA and PI double staining in a time course (Figure 2-13A). The results showed that this period for loss of FDA signal and increased PI signal was longer in the triple mutant (Figure 2-13B). ASPA2 overexpression lines shortened this time period and catalytic inactive protease overexpression lines did not show change this time

period (Figure 2-13B). This means that it takes a longer time for the triple mutant cells to exhibit an increase in cell membrane permeability. This suggests that when these proteases are insufficient, the digestion of membrane components is slower, and the membrane system remains intact and ordered for a longer time in the mutant cells. Reports showed that the lateral root caps serve as an auxin sink. Disturbing PCD in lateral root caps affect the auxin distribution along the roots and thus affect lateral root formation (Xuan et al., 2016). With sufficient nitrogen supply, the triple mutant roots showed more lateral roots (Figure 2-13C). This result indicates that ASPA2 may regulate root morphology through regulating PCD in the root.

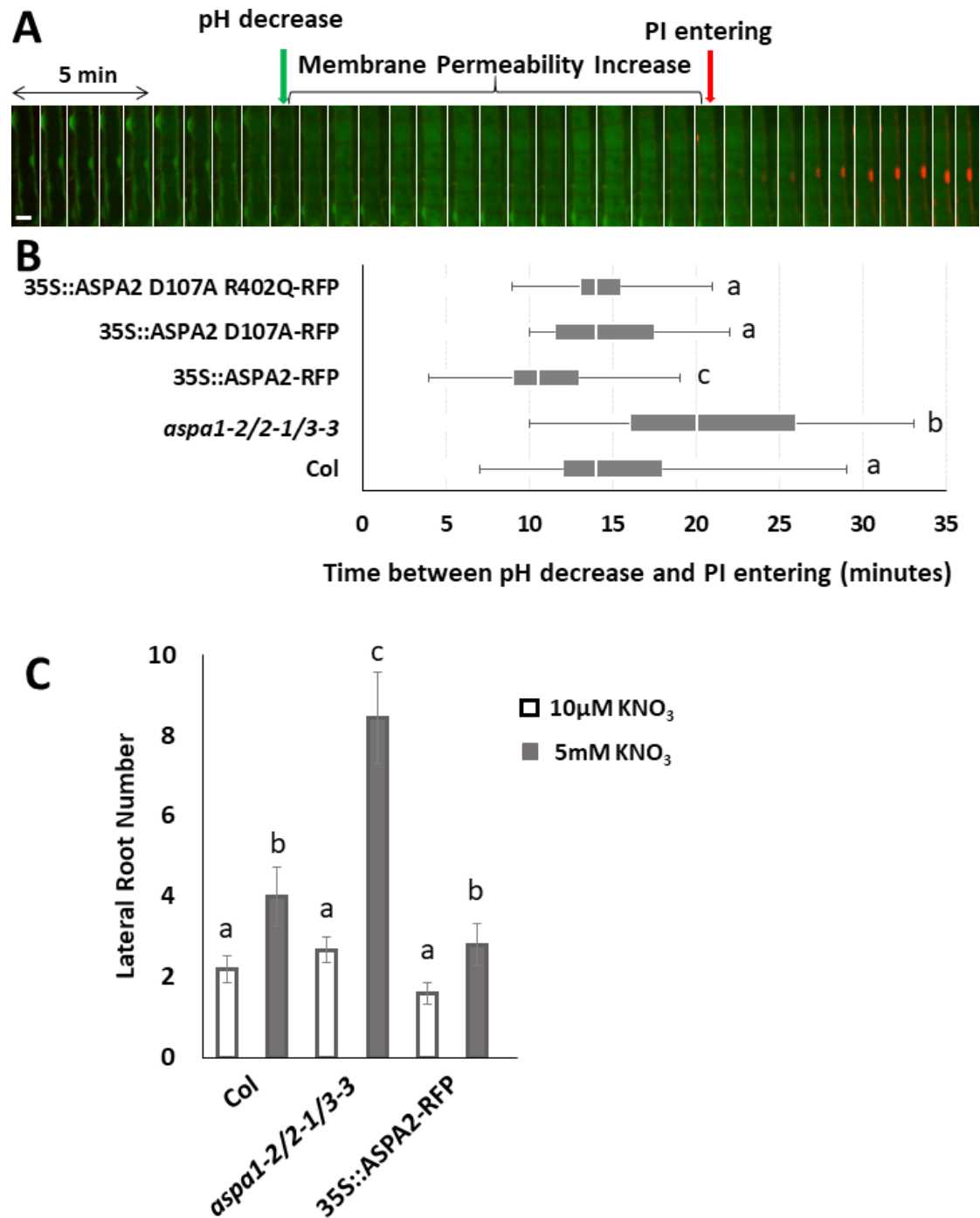


Figure 2-13. Fluorescent diacetate (FDA) and propidium iodide (PI) double staining in lateral root cap cells in Col and *ASPA* mutants over time. (A) Representative image of time course of fluorescein diacetate (FDA) and propidium iodide (PI) double staining in 6DAG Col seedling lateral root cap. Bar = 20μm. (B) Time between disappearing of FDA signal and appearing of PI signal in Col-0, *aspa1-2/2-1/3-3*, 35S::ASPA2-RFP, 35S::

ASPA2 D107A-RFP, 35S:: ASPA2 D107A R402Q-RFP in lateral root caps. N=13-30 cells for each line. Different letters indicate statistical significance, $p < 0.05$ ANOVA followed by Tukey post-hoc test. (C) Lateral root number in Col-0, *aspa1-2/2-1/3-3* and 35S::ASPA2-RFP under different nitrogen conditions. seedlings were grown on 1/4MS media for 4 days and then transferred to media with low nitrogen (10 μ M KNO₃) or sufficient nitrogen (5mM KNO₃) for additional 4 days. Different letters indicate statistical significance, $p < 0.05$ ANOVA followed by Tukey post-hoc test.

Discussion

Aspartic proteases (ASPAs) in seed development and germination

Aspartic proteases have long been believed to function in seed maturation due to their high expression levels in seeds in several plant species, such as cardoon and barley (Pereira et al., 2008; Sarkkinen et al., 1992). They have also been found to function in both seed maturation and seed germination processes (Wrobel et al., 1992). However, due to the lack of loss-of-function mutants, the role of aspartic proteases during these two processes *in vivo* was still unclear. To test this hypothesis, by molecular genetic study by using the *Arabidopsis* mutants, this dissertation shows that insufficient aspartic protease activity led to delayed seed maturation and delayed germination. *ASPA1* was proposed to first function in MVB for seed storage processing at torpedo stage (Otegui et al., 2006), and *ASPA1* was frequently chosen as the marker in seed development as well. *ASPA1* may have other targets such as other proenzymes,

but so far there are no further reports on this. Without enough ASPA proteases, the accumulation of seed storage proteins slows down, and the developing seeds show delay desiccation. The extended time for seed storage protein accumulation compensates for this and thus the total amount of storage proteins is higher. This might be the reason why the mutant seeds were larger and weighed near twice as the wild type.

Seed storage protein processing may not be the only function in seed maturation. Some proteases may not be active or not directly involved in proteolytic activity. They may be packaged during seed dormancy, then are activated during seed imbibition and germination. This possibility for the ASPAs could be inferred by the impact on seed germination in the mutant seeds. Though there is indeed *ASPA2* expression during germination, the major source of proteases in imbibed seeds are likely to be the ones stored during seed maturation. The reason is that if most proteases were newly synthesized, gibberellin treatment should have rescued the phenotype to some extent as other proteases may compensate for some of the missing ASPAs. The major function of ASPAs during seed germination is likely to be degradation of seed storage proteins for the growing young seedling.

On the other hand, the delayed fusion of seed storage vacuoles may suggest its second role during seed germination in membrane disturbance. This comes from the structural aspect that the PSI in aspartic proteases is quite unique from other proteases. Besides the fact that this PSI is cleaved from the protease, there is a

hypothesis that this PSI may function as an independent protein in membrane disturbance. The logic supports this hypothesis is that in MVBs, there are smaller compartments surrounded by intact membranes, and therefore the contents inside these compartments are not released for degradation. A protein that disrupts the membrane structure under low pH may help the proteases interact with their targets. And aspartic proteases are good candidates for these functions. The vacuole fusion during seed germination supports the role of membrane disturbance for aspartic proteases. However, this result did not show whether this function comes from the PSI or the proteolytic domains.

Thus, ASPAs are important for plant seed development and germination. Delayed seed development and germination are not advantageous traits in evolution, because the plants may not respond to the proper time for seed maturation and seedling growth.

ASPAs function in tissues other than seeds

Though ASPAs were first found in seeds, *ASPA1* and *ASPA2* were expressed almost throughout the whole plant, but their functions in these tissues haven't been reported yet. The role in seed maturation and germination suggests the hypothesis that ASPAs process the bulk of the targets and regulate nitrogen supplies by proteolytic processing of aged or broken proteins in the cell. The reduced response to low nitrogen in both mutant seedlings and overexpression plants support this hypothesis. In both seed

maturation and low nitrogen conditions, ABA is an important signaling component. The transcriptional results show that *ASPA2* expression was slightly upregulated to low nitrogen, and highly upregulated by ABA. *ASPA2* promoter region contains ABI binding elements. In contrast, *ASPA1* do not have these elements in the promoter, and the expression level remained at a constant level. This is consistent with the reported results that *ASPA1* expression is relative stable (Endo et al., 2014). The differences between *ASPA1* and *ASPA2* indicate that *ASPA1* is more likely a housekeeping gene and *ASPA2* is responsive to environment stresses. Another ABA regulated physiological process is stomata opening and drought tolerance. *ASPA1* overexpression has been reported to enhance drought tolerance in *Arabidopsis* by regulate stomata opening (Sebastián et al, 2020). However, this is more likely to mimic another aspartic protease *ASPG1* (*ASPARTIC PROTEASE IN GUARD CELL1*) function in guard cells which does not contain a saposin-like domain (Yao et al., 2012).

The plant specific insert is important for the aspartic protease vacuolar targeting, and this is also the case for *ASPA2*. The vacuolar targeting of the catalytic inactive form of *ASPA2* suggests that the self-catalytic activity is not required for vacuolar targeting either. The processing of *ASPA2* is likely to occur via other proteases. The primary function of PSI is interacting with lipids, and PSI from potato StAP shows the lipid interaction activity *in vitro*. The newly identified six amino acids motif in potato StAP PSI shows its important role in conformation change. Abolishment of this motif blocks the conformation change and blocks interactions with lipids. As a result, it was

hypothesized that mutation in this motif in ASPA2 PSI would also block conformation change, and thus block its function in the vacuolar targeting. However, the results showed that the mutated ASPA2 was still trafficked to the vacuole. This result suggests that this motif in PSI is not responsible for vacuolar targeting. It could be possible that the PSI interaction with lipid membranes doesn't require conformational change in *Arabidopsis* for vacuolar targeting, or there is another novel mechanism for PSI function *in vivo*. The catalytically inactive form of ASPA2 (ASPA2-D107A) could be regarded as a "native" version of PSI. In this dissertation, no significant phenotype was observed in this mutated ASPA2 overexpression lines. One possible reason is that PSI is not directly involved in the normal plant growth. Potato PSI form StAP show anti-bacteria activity (Muñoz et al., 2010; Fery et al., 2018). Overexpression potato PSI in *Arabidopsis* led to enhanced resistance to *Botrytis cinerea*, and the plants were taller than wild type (Frey et al., 2018). Overexpression of the catalytically inactive ASPA2 D107A in *Arabidopsis*, the plants did not show a higher height in this dissertation.

This anti-bacterial activity for PSI requires that the PSI is secreted to the extracellular space. However, no reports have been shown that PSI is able to traffic to the extracellular space. The artificial recombinant PSI constructs show that PSI still traffics to the vacuole (Vieira et al., 2019). However, it is possible that under certain circumstances, PSI is secreted in plant defense responses, which needs further studies.

In terms of plant defense, another interesting thing is, the mature ASPAs (which does not have PSI) show structural similarity with the anti-pathogen protein xylanase

inhibitor 1 from wheat (Fierens et al., 2003; Sansen et al., 2004). Xylanase is synthesized and secreted from the pathogens and digest plant cell walls to attack the plants. Xylanase inhibitor 1 binds and inactivate xylanases (Fierens et al., 2003). Xylanase inhibitor 1 lacks essential catalytical residuals and is proteolytically nonfunctional (Sansen et al., 2004). This provides another possible function of ASPAs in plant defense. If ASPAs are secreted to the extracellular space, they might have the ability to inactive xylanases, and this function is independent of PSI. It is interesting that a single peptide encodes two independent functional units functioning in plant defense. Further studies will explore whether the ASPAs are able to traffic to extracellular space in plant defense responses.

ASPAs in programmed cell death (PCD)

The properties of ASPAs *in vitro* is well-studied. However, from these studies, these proteins seem to be simply a tool for proteolytic activity without any specificity. If this is the case, it seems that PSI is only necessary for vacuolar targeting. However, the PSI structural feature is conservative in plants, which presupposes to a function that is important yet known in plant growth and development. The *ASP3* expression pattern seems to provide hints on their functions. The roles of *ASP3* in PCD has long been proposed due to its restricted expression pattern. But the lack of PCD-related phenotypes in single knockout mutant makes the exact role remained unclear. This may partially result from the redundancy of *ASP1* and *ASP2* in those tissues. This

leads to another hypothesis that ASPAs function in regulating PCD in *Arabidopsis*.

In PCD cells, a set of proteins are usually co-expressed such as CEP1, ASPA3 and BFN1. These proteins function in the last stages of PCD for nutrient recycling and cell components disruption. While most upstream transcriptional factors show tissue specific expression pattern, the set of CEP1, ASPA3 and BFN1 is expressed in almost all the PCD tissues. But the expression time is different for these genes. In *Arabidopsis* stigmas, the expression order is *CEP1* first, then *ASPA3*, and *BFN1* is the last (Gao et al., 2018). CEP1 functions in the cytosol, and it is likely to participate in the signaling transduction. BFN1 functions in the nuclei for the final degradation of DNA. ASPA3 may be the primary protease that is involved in bulk proteolytic activity of proteins for recycling nitrogen for other tissues. The membrane disturbance ability is also an advantage during this process. The results here showed that insufficient ASPAs reduced the rate of membrane permeability increase in lateral root cap, and delayed xylem maturation in the root. A delay in PCD processes may impact on plant growth and development, such as root architecture since PCD occurs in root caps, tracheary elements and the base of lateral roots. For example, lateral root caps are believed to be an auxin sink and the peeling off affects the release of auxin. Therefore, the position of lateral root cap PCD affects the auxin distribution in the root tip, and thus affect the auxin distribution along the root (Xuan et al., 2016). The distribution of auxin, or the maximum auxin sites along the root, is associated with lateral root formation (Wei et al., 2016). As a result, PCD is associated with root architecture in *Arabidopsis*. ASPAs

regulate the rate of PCD in lateral root caps and affect root architecture.

Conclusion

In summary, the biological functions of ASPAs might be bulk proteolytic activity in vacuoles for nitrogen recycling. The membrane disturbance activity promotes interaction between proteases and substrates. *ASPA1* appears to be a housekeeping protease and *ASPA2* functions in response to environmental stresses. These two proteases are expressed in most plant tissues. *ASPA3* functions in PCD tissues as an additional contributor to the degradation events. ASPAs are involved in seed development and germination, as well as programmed cell death in the plants. The independent function of PSI was not found in this dissertation.

There are still some unresolved questions. First, the triple mutants do not show a dramatic phenotype in older seedlings or adult plants (Figure S2 and S3). This may be due to compensation from a low level of *ASPA1* activity in the triple mutants. As *ASPA1* expression is normally relatively high throughout the plants. The remaining protease activity may still be enough for the basic metabolic requirements. As a result, the *aspa1* knockout mutant is preferred for studying the biological functions of ASPAs *in vivo*. Generating the *aspa1* knockout mutant by CRISPR is a good choice. One of the future directions is to create a knockout triple mutant for phenotypic studies and compensating this knockout mutant to determine the function of these proteins *in vivo*.

Second, the biological role of PSI other than vacuolar targeting remains unclear. Overexpression of the catalytic inactive ASPA2 D107A in *Arabidopsis* did not affect plant growth. The root growth and PCD in lateral root caps were not affected either in ASPA2 D107A overexpression lines. These results suggest that PSI does not function independently from the proteolytic domain. PSI overexpression does not promote PCD in lateral root caps (Figure 11) or enhance plant growth like StAP PSI (Figure S2 and S3). The major function of PSI is associating the protease domain with membranes, bringing it to vacuoles. Glycosylation may be a signal for ASPA transport to TGN so that vesicles containing ASPAs could fuse with early endosomes for plasma membrane protein digestion. One of the future aspects is to find whether ASPAs or PSIs are secreted to the extracellular space. This will provide information on whether they may be involved in plant defense response.

Third, the role of ASPAs in programmed cell death needs further studies. The difference between wild type and the triple mutant was subtle, and the only phenotypes found were in tracheary element maturation and lateral root cap turnover. In other PCD tissues such as the tapetum, PCD related phenotypes were not detected. The knockout triple mutants will also help with this question. Another direction is to further mutate co-expressed genes such as *CEP1* and *BFN1* and explore how these genes affect PCD in combination.

This dissertation demonstrated the role of ASPAs processing seed storage proteins in seed germination *in vivo* for the first time. And this dissertation also

provides evidence of the involvement of ASPAs in programmed cell death by promoting membrane permeability. These results broaden the knowledge of the multiple roles of aspartic proteases in plant growth and development.

Materials and Methods

Plant materials

All the *Arabidopsis thaliana* plants are in the Columbia-0 (Col-0) ecotype genetic background. T-DNA insertional mutants (*aspa2-1* SALK097505; *aspa2-2* SALK021601; *aspa1-1* SALK092586; *aspa1-2* SALK041027; *aspa3-3* SALK056711) were sourced from The Arabidopsis Biological Resource Center, The Ohio State University (ABRC; www.abrc.osu.edu). T-DNA insertions were confirmed by PCR with the primer in T-DNA sequence (LBb1.3) and the primer in the flanking gene regions (primer RP). The primers used in genotyping are listed in Table S01. The expression level of each ASPA was detected by real-time PCR and the primers for real-time PCR are listed in Table S01. Details on methods for genotyping and real-time PCR are described in appendix D. For germination on solid media, *Arabidopsis* seeds were surface sterilized by soaking in 20% bleach (containing sodium hypochlorite) for 15 minutes with agitation. Seeds were then rinsed 3-5 times in sterile water. Seeds were sown on 1/4 Murashige and Skoog (MS) medium (RPI Corp.) media containing 0.5% sucrose with 0.8% agar. Seeds were stratified at 4°C for 2 days in the dark and then placed in growth chamber at 22°C, with 24 hr continuous white light at 100 $\mu\text{mol m}^{-2} \text{s}^{-1}$. For chemical treatment,

seedlings were first grown for 4 days on regular 1/4MS media, then transferred to new media containing the corresponding chemicals. The working concentrations of chemicals used in this research were: 1 μ M Gibberellin acid; 10mM hydrogen peroxide (H₂O₂); 10 μ M diphenyleneiodonium (DPI); 10 μ M brefeldin A (BFA) ; 100nM concanamycin A (conc A); 4 μ M propidium iodide (PI); 5 μ g/ml fluorescein diacetate (FDA); 2 μ M abscisic acid (ABA); 75mM sodium chloride (NaCl). Solvent (ethanol or DMSO) was added as the control, and the concentration was the same with the corresponding chemical concentration in each experiment. Low nitrogen media was prepared by adding 10 μ M potassium nitrate (KNO₃) in 1/4MS without nitrogen (MS w/o nitrogen) media. Sufficient nitrogen media was prepared by adding 5mM KNO₃ in 1/4MS w/o nitrogen media.

Adult plants were grown in growth chamber at 22°C with 16 hr light and 8 hr darkness cycles. Light intensity was 100 μ mol m⁻² s⁻¹ with a mixture of fluorescent and incandescent bulbs. Relative humidity was 50%. For seed weight measurement, seeds were harvested and stored in drying chamber containing drierite for at least three days.

Germination test

Seeds were harvested and stored in drying chamber for at least three days. Only seeds harvested within a month were used. Each time, 3 biological replicates were measured. Each replicate contained around 150 seeds. Seeds were sterilized and sowed as mentioned above. Seeds were placed at 22°C, with 24 hr continuous light at

60 $\mu\text{mol m}^{-2} \text{ s}^{-1}$. Germinated seeds were counted every 12 hours. Germination was defined as radicle emergence.

Seed Protein extraction

For each time point, 200 *Arabidopsis* seeds were accurately counted and imbibed in sterilized water. Seeds were ground with a grind stick in Eppendorf tubes on ice. The ground seeds were immediately resuspended in 100 μL protein extraction buffer (50 mM sodium citrate, pH 5.5; 5% SDS (w/v); 0.01% BSA (w/v); 150 mM NaCl; 2% (v/v) β -mercaptoethanol and 1 μL of protease inhibitor cocktail (Genesee Scientific). The mixture was incubated for 60 minutes at 100° C. Samples were centrifuged at 4° C, 14,000g for 30 minutes and the supernatant was collected. The samples were stored in -80° C if not used immediately.

SDS-PAGE

Total proteins were separated by SDS polyacrylamide gel electrophoresis (SDS-PAGE). 10 μL samples were prepared by adding 2 μL of 6X SDS (sodium dodecyl sulfate) loading buffer (1.2g SDS, 0.01% bromophenol blue, 4.7ml glycerol, 1.2ml Tris 0.5M pH=6.8, 2.1ml ddH₂O). Samples were loaded onto 12% polyacrylamide 0.75mm 10-well or 15-well gel (Bio-Rad®). Precision Plus Protein Dual Color Standards (Bio Rad) was used as marker size. Electrophoresis was carried out in 1X Running Buffer (3g of Tris base, 14.4g of glycine, and 1g of SDS in 1000 ml water) at 120V for approximately 4 hours or until the dye front reached the front of the gel. Seed proteins were

visualized by staining with Coomassie Blue.

Coomassie blue staining

For visualization of seed storage proteins, the gel was stained by incubating overnight in 20ml Coomassie staining solution (0.1% Coomassie bright blue in 50% methanol, 10% acetic acid). The gel was de-stained for 3 hours with de-staining solution (10% acetic acid, 50% methanol) with at least two changes of this solution until the background was nearly clear.

Glycosylation test

300mg *Arabidopsis* leaves were harvested and then ground with a grind stick in the Eppendorf tube with liquid nitrogen. The ground tissues were resuspended in 300 μ L protein extraction buffer, incubated for 60 minutes at 100° C, centrifuged at 4° C, 14,500g for 30 minutes and the supernatant was collected.

Glycosylation was detected by Endo Hf (New England BioLabs) digestion according to the manufacturer's instruction. Briefly, 17 μ L of the extracted protein sample was added with 2 μ L 10xGlycoBuffer 3 and 1 μ L Endo Hf. The sample was incubated at 37°C for 1 hour. Then the sample was analyzed by SDS-PAGE and Western blot.

Western blot

For immunoblotting, proteins were transferred to polyvinylidene difluoride (PVDF)

membrane in Tris-glycine-methanol transfer buffer (2.9g glycine, 5.8g Tris, 0.37g SDS 100mL methanol, 900mL water) at 120V for 80 minutes at 4°C and then rinsed briefly in 1xPBS. Membranes were blocked overnight at 4°C in blocking buffer (5% non-fat milk in 1xPBS with 0.02% Tween20) or 1.5 hours at room temperature. The membrane was rinsed gently with washing buffer (1% non-fat milk in 1x PBS with 0.02% Tween20) for three times, 15 minutes each. The membrane then was incubated with primary antibody (anti-HA) in blocking buffer overnight at 4°C or 1.5 hours at room temperature. The membrane was rinsed with washing buffer for three times and each time for 15 minutes. Then the membrane was incubated with secondary antibody (anti-rabbit digoxigenin) at room temperature for 1.5 hours. The membrane was rinsed with washing buffer for three times and each time for 15 minutes. Proteins were visualized using a SuperSignal West Femto Kit (Thermo Scientific). Images were taken by C-DiGit Blot Scanner (LI-COR).

Cloning and expression vector construction

ASPA1, *ASPA2* and *ASPA3* coding sequences (CDS) were cloned without the stop codon from seedling cDNA. Fragments were inserted into pDONR/Zeo by BP reaction. The entry clones were confirmed by sequencing. Primers for cloning, site-direct mutagenesis and sequencing are listed in Table S01. The entry clones were incorporated into pH7RGW2 (35S promoter, RFP tag fused on C-terminus) and pEarleyGate102 (35S promoter, CFP and HA tag fused on C-terminus) by LR reactions.

The expression constructs were confirmed by sequencing and the corresponding primers are listed in Table S01. Transgenic plants were created by floral dipping method (Clough and Bent, 1998) with agrobacterium strain GV3101.

For promoter fusions with the histone tag, first the promoters of *ASPA1* (1.9kb), *ASPA2* (1.9kb) and *ASPA3* (2.0kb) were cloned and inserted into pGEM-T-Easy vector. Entry clones were confirmed by sequencing. Primers for cloning and sequencing are listed in Table S01. Gateway vector pUBC::RFP-Dest (Grefen et al., 2010) was digested by *SpeI* and *PsiI* and ligated with mCherry. Then the modified pUBC::mCherry vector was digested by *SacI* and *PspXI* and ligated with *ASPA1* promoter. The gateway vector pUBC::YFP-Dest was digested by *SacI* and *PspXI* and ligated with *ASPA2* promoter. The gateway vector pUBN::YFP-Dest was digested by *SacI* and *PspXI* and ligated with *ASPA3* promoter. Histone 2A 10 CDS was cloned from seedling cDNA with and without the stop codon and inserted into pDONR/Zeo vector. The final constructs *ASPA1* promoter::H2A-mCherry and *ASPA2* promoter::H2A-YFP were created by LR reaction. The expression constructs were confirmed by sequencing and the corresponding primers are listed in Table S01. Details on methods of molecular cloning are described in appendix D.

Microscopy

Confocal microscopy was carried out using a Zeiss LSM 710 Confocal laser scanning microscope (Carl Zeiss, Germany) with Axio Imager 2. Pixel dwell time was

0.01 ms. The master gain was always set to less than 893, with a digital gain of 1.5. For RFP/mCherry acquisition: 594 nm (5%) excitation and 588-696 nm emission. For YFP acquisition: 514 nm (5%) excitation and 519-560 nm emission. For GFP: 488 nm (5%) excitation and 493-598 nm emission. For CFP: 458 nm (5%) excitation and 453-580 nm emission. For PI: 543 nm (5%) excitation and 583-718 nm emission. For FDA: 488 nm (5%) excitation and 493-583 nm emission. Quantification of fluorescence intensity was analyzed using ZEN Lite 2012. All images were processed with ZEN Lite 2012 (Zeiss) and ImageJ.

Time-course image of PI (propidium iodide) and PI/FDA (fluorescein diacetate) double staining in lateral root cap

To keep the roots alive for a long time while imaging, seedlings were imaged in the 35 mm petri dish with high precision 1.5 coverslip on the bottom (14 mm glass diameter, MatTek). Seedlings were placed with 1/4MS (0.5% sucrose, 1% agar) containing PI only (4 μ M) or PI (4 μ M) and FDA (5 μ g/ml). Images were taken by every minute or every five minutes as noted. All images were processed with ZEN Lite 2012 (Zeiss) and ImageJ.

Chapter 3: Elucidating features and functions of plant prosaposin-like proteins (PSAPLIPs)

Abstract

Saposin-like proteins have been well studied in animals. In plants, the only reported saposin-like structure is the plant specific insert in some aspartic proteases. Another type of saposin-like (SapB-like) proteins, the prosaposin-like proteins (PSAPLIP) have been paid less attention. These proteins are ubiquitously present across the plant kingdom from green algae to flowering plants, indicating their importance in plant growth and development. Here alignment and comparison of protein sequences among different species revealed the high similarity between plant prosaposin-like proteins and human prosaposin in primary and secondary structure. Unique features of prosaposin-like proteins in plants were also identified. PSAPLIPs contain two SapB-like domains in angiosperms, while in gymnosperm, moss, liverwort and green algae, most PSAPLIPs contain three SapB-like domains. In most species, there are 1-4 PSAPLIPs encoded genes in their genomes, and *Arabidopsis thaliana* has two PSAPLIPs, *AtPSAPLIP1* (At3g51730) and *AtPSAPLIP2* (At5g01800). Both *AtPSAPLIP1* and *AtPSAPLIP2* were targeted to vacuoles and both proteins were sensitive to concanamycin A treatment. However, *AtPSAPLIP1* was sensitive to brefeldin A treatment while *AtPSAPLIP2* was not. The promoter reporter activity results showed

that *AtPSAPLIP1* was primarily expressed in sepals and pollen grains, while *AtPSAPLIP2* was expressed in petals and young anthers. These results suggest the important role of prosaposin-like proteins in reproductive organ development, especially in male gametophyte development. They may facilitate degradation of target signaling proteins in the cell. This dissertation characterized the plant prosaposin-like proteins for the first time and provided insights on a new class of proteins regulating male gametophyte development in plant reproductive process.

Introduction

Saposin-like proteins (SAPLIP) are named after saposins, which contain four small proteins derived from one single precursor called prosaposin. Saposins are important in cellular metabolism as cofactors in sphingolipid catabolism (Bruhn, 2005). SAPLIPs are found throughout eukaryotes from amoebozoans to mammals. SAPLIPs exhibit low sequence similarities among different species, but they are conserved in the six conserved cysteines and several conserved hydrophobic and polar charged residues which enable the protein folding into the conformation to interact with lipids (Bruhn, 2005). In animals, SAPLIPs are found to participate in a variety of different pathways, such as co-factors of lipid-degrading enzymes (Kishimoto et al., 1992; Schuette et al., 2001), surface tension regulator (surfactant protein B) (Cochrane et al., 1991), antimicrobial effector (Pena et al., 1997). Some SAPLIPs activities are independent of lipid interactions, such as J3 crystallin found in jellyfish *Tripedalia cystophora*

(Piatigorsky et al., 1997). Loss of SAPLIP function in mammals is associated with diseases states, such as deficiency in saposin C leading to Gaucher disease which is a type of lysosomal storage disorder (Tamargo et al., 2012), and deficiency in saposin A leading to Krabbe disease which is a disorder that the protective coat in nerve system is defective (Spiegel et al., 2005).

Structural similarity is high among SAPLIPs. There two types of conformation reported, and they show slightly different. NK-lysin is one type and the SAPLIP domain contains 5 helices fold into two halves. The first half consists of helices 4 and 5 packed perpendicularly against helix 1. The other half contains helix 2 and 3 (Liepinsh et al. 1997). Saposin B is representative of other types of SAPLIPs. The two halves of saposin B crystallizes as a dimer into a shell shape. The saposin B monomer has four helices and shows an open formation in a V shape. This has been proposed as the lipid binding position (Ahn et al., 2003).

Both types of conformation are in favor of lipid interaction. The soluble, monomeric form of SAPLIP holds a closed conformation with the hydrophobic surface hidden in the cavity. Charged residues mediated the initial contact with the negatively charged lipid membrane surface by electrostatic interactions. Then the protein change into open conformation, probably associated with dimerization or oligomerization. The membrane-embedded oligomer is hypothesized to form a pore in the membrane allowing presentation to the hydrolytic enzymes (Rossmann et al., 2008; Olmeda et al., 2012). And in the end, either the lipids are extracted or two adjacent membranes fused

by saposin-like proteins.

Animal SAPLIPs have been extensively studied on the structures and functions (Azuma et al., 1994; Ciaffoni et al., 2001; Ahn et al., 2003; De Alba et al., 2003; Kang et al., 2004; Hawkins et al., 2005; Hill et al., 2006; Popovic et al., 2008; Olmeda et al., 2013). However, plant SAPLIPs are not studied as extensively as animals. Plant SAPLIPs are generally referred as a domain called plant specific inserts (PSI) in aspartic proteases (Brodelius et al., 2005). In general, plant PSIs are similar with human saposins in terms of sequence features and the overall structure. The PSI from phytepsin in barley shows highly structural similarity with NK-lysin (Kervinen et al., 1999). PSI from cardoon also shows high similarity to human saposins C and it is able to activate human glucosylceramidase *in vitro* (Brodelius et al., 2005). PSI of StAP in potato is able to induce vesicle disruption *in vitro*, similar with human saposin C and the secondary conformation is pH-dependent which is similar to human saposins (Bryksa et al., 2011).

Sequences are highly similar and conserved among plant PSIs (Bryksa et al., 2017). They all exhibit leakage activity in bilayer composed of a vacuole-like phospholipid mixture and membrane fusion activity *in vitro*. This activity is pH-dependent and the optimal pH is 4.5 and requires the presence of acidic phospholipids such as phosphatidylserine. Low pH results in dimerization of potato PSI, and the monomer is prevalent under neutral pH. All these behaviours are similar to mammalian saposins

Although there are a lot of similarities between plant PSI and mammalian saposin-

like proteins, there are some features unique to plants. A recent study showed that conformation change is the molecular basis of bilayer membrane leakage at low pH. A novel six-residue motif in H3 helix ([N/Q]-[N/Q]-[N/Q]-[A/L/I/V]-[K/R]-[N/Q]) was identified which accounts for this configuration change. A point mutation K83Q in this motif in helix H3 blocks the response to low pH activation with respect to conformation change (Bryksa et al., 2017). This motif may be responsible for lipid-interactions as this motif is also found in several other membrane-interacting proteins in different plant species (Bryksa et al., 2017). But This motif is not seen in human and other mammalian saposins. Another difference between PSI and mammalian saposin is that the orientation of helices is switched from N terminus to C terminus (Bliven et al., 2012). The overall configuration of the secondary structure is not affected. This could be evolved from gene duplication and subsequent deletion event in plant evolution history. The third difference is that the PSI in aspartic proteases is cleaved off to produce the mature protease, while in mammals, saposin-like domains are still linked to the mature proteins (Bruhn, 2005).

PSI is important for the aspartic protease vacuolar targeting *in vivo* (Kervinen et al., 1999; Terauchi et al., 2006). Overexpression of PSI from the potato aspartic protease in *Arabidopsis* enhances the plant resistance against *Botrytis cinerea* (Frey et al., 2018). However, the independent function of PSI *in vivo* still lacks experimental supports. This leads to one hypothesis: there are other SAPLIPs in the plants which are not characterized yet.

With information from Uniprot and other protein databases, there are a group of uncharacterized saposin-like domain containing proteins in plant genomes. Almost no reports studied on the structural features or functional analysis of these proteins in the plants, so little information is available. In this dissertation, the primary and secondary structures of these proteins from across the plant kingdom were predicted and analyzed to understand the distribution and diversity of these proteins.

By sequence screening with the keyword search in Uniprot, sequence alignments and comparisons, this group of proteins showed a high similarity with the human prosaposin. As a result, they are named prosaposin-like proteins (PSAPLIPs) in this dissertation. This analysis showed that PSAPLIPs are found in all plant phyla and the number of genes varies. With structural prediction in Phyre2, saposin-like domain from *Arabidopsis* PSAPLIPs showed high similarity to human saposins. *Arabidopsis AtPSAPLIP1* and *AtPSAPLIP2* were analyzed spatiotemporal expression and subcellular targeting. The results suggest that PSAPLIPs are important in male gametophyte development in *Arabidopsis*, possibly by facilitating target signaling proteins trafficking to the vacuole for degradation.

Results

Phylogenetic studies of PSAPLIPs in plants

More than 160 PSAPLIP genes from 67 species have been identified in higher plants via pairwise ortholog predictions (EggNOG, eggnogdb.embl.de). In Uniprot

database, more than 2000 proteins are annotated as containing the saposin-like domains. Some of them are aspartic proteases, which contains the saposin-like domain called the plant specific insert (PSI). The remaining 459 proteins were uncharacterized in 152 plant species from green algae to flowering plants.

In angiosperms, there are 417 sequences annotated as containing at least one saposin B like domain (SapB-like domain). These sequences can be divided into two groups depending on their predicted sequence structures. The first one contains the N-terminal signal peptide and SapB-like domain of 80-82 amino acid residues, which is the typical length of a saposin-like protein. The second group contains an N-terminal signal peptide, an annotated saposin-like domain around 130 amino acid residues, following by disordered regions in C terminus, usually with a polyampholyte region. From amino acid alignments, these sequences have clearly diverged from the other SapB-like domain containing proteins. The sequence features of typical SAPLIPs, such as the distribution of the six conserved cysteines, and conserved hydrophobic and polar residues, were not found in this group of proteins. This second group is more similar with human nucleophosmin than to saposins and this group is classified as nucleophosmin family by gene ontology (PANTHER, pantherdb.org). The reason that members of this group were predicted to be saposin-like proteins is likely due to the prediction algorithm is based on the six conserved cysteines in saposin-like proteins, and plant nucleophosmins happen to contain several cysteines in their sequences. As a result, these sequences are likely incorrectly auto-predicted by Uniprot database,

and this group should be considered and annotated as nucleophosmin-like proteins rather than saposin-like proteins. Among the 417 uncharacterized angiosperm sequences, 73 belongs to nucleophosmin family. The remaining 344 should be considered the PSAPLIPs in plants. Then the sequences annotated as incomplete were also excluded. The remaining sequences were used for the following analyses.

Typical PSAPLIPs contain an N-terminal signal peptide and two saposin-B like domains. PSAPLIPs are predicted to be in the vacuole. However, possibility of secretion to extracellular space or other compartments could not be excluded. The *Arabidopsis AtPSAPLIP1* is annotated in both the cytosol and the vacuole.

Several proteins have no N-terminal signal peptide (SP) prediction. In most sequences, the signal peptide is around 18-35 amino acid residues. In those gene apparently missing SP sequences, there are indeed sequences long enough to be a signal peptide. Novel types of signal peptide may exist in these species. Only six sequences show an N-terminal domain less than 10 amino acid residues (*Ananas comosus* ACMD2_06262; *Arundo donax* no gene ID, protein ID A0A0A9V254; *Dichanthelium oligosanthes* BAE44_0009052; *Panicum miliaceum* C2845_PM06G26640; *Helianthus annuus* HannXRQ_Chr10g0286291; *Dorcoceras hygrometricum* F511_29468) (Figure S08). However, these annotations were auto-predicted from the genomic DNA sequence and It is possible that the start codon was predicted incorrectly. If these genes indeed lack a signal peptide, this suggests that these PSAPLIPs may either function in other cellular compartments or traffic to vacuole

by other facilitating proteins.

Human prosaposin is processed into four mature saposins, and human saposin B remains dimerizes under most conditions (Hiraiwa et al., 1993; Kishimoto et al., 1992; Leonova et al., 1996). It can be inferred that Sap-B like domains may also function as dimers, and this may be the reason that most plant PSAPLIPs contain two Sap-B like domains. In some genes, only one Sap-B domain is found (36 sequences among all 344 angiosperm PSAPLIPs), but it is more likely a prediction error because all these predictions are from genomic DNA sequence. Some sequences appear to lack the N-terminal part of the SapB-like domains and some seem to lack C-terminal part. This suggests that they may be fragments but incorrectly annotated as complete ones, or they may be complete sequences. This would need to be confirmed in these species by further studies. Overall, most identified angiosperm PSAPLIPs contains two SapB-like domains, and it can be concluded that this is the prevalent form in angiosperms (Figure 3-02).

In green algae, liverworts, mosses and gymnosperms, the copy number of SapB-like domain varies. The gymnosperm, *Picea sitchensis* contains a protein with two SapB-like domains (Uniprot protein ID A9P228) and a protein with three SapB-like domains (Uniprot protein ID A9P283). While PSAPLIPs from *Araucaria cunninghamii* (Uniprot protein ID AOAOD6R2G) and *Wollemia nobilis* (Uniprot protein ID A0AOC9RXJ5) contain three copies of Sap-B like domains in PSAPLIPs and no PSAPLIP was identified with two SapB-like domains in these two species. Due to the limited

data for ferns, no PSAPLIPs were found in ferns. In liverworts and mosses, only three PSAPLIPs were found and they contain three SapB like domains (*Chara braunii* CBR_g3540; *Physcomitrella patens subsp. patens* PHYPA_022478; *Physcomitrella patens subsp. patens* PHYPA_018982). In green algae PSAPLIPs, the number of SapB-like domains varies from one to three. However, most sequences are derived from whole genome shotgun (WGS) entries (an EMBL/GenBank/DDBJ), therefore this is an initial analysis based on preliminary data. Three SapB-like domains still appear to be the major form of PSAPLIPs in green algae (Figure S09).

The trend of evolution of plant PSAPLIPs can be depicted as following: in green algae, genes with three SapB-like domains is the prevalent form. In liverworts and mosses, only genes with three SapB-like domains are found. This supports the hypothesis that land plants evolved from single origin that contained only this type of *PSAPLIP* genes. In gymnosperms *Picea sitchensis*, genes with two and three SapB-like domains are found. This suggests that genes with two SapB-like domains first evolved in gymnosperms, and probably evolved from deletion of one of the SapB-like domains after a gene duplication event. In angiosperms, no *PSAPLIP* genes were found with three SapB-like domains. *PSAPLIPs* with two SapB-like domains are found in all angiosperms reported in Uniprot. This supported the single origin for all angiosperms. *PSAPLIPs* with only one SapB-like domain may exist in some angiosperm species, like the grape, but this hypothesis is not supported by current information. Sequence data from more species, and more experimental sequence data are needed to obtain a

better understanding of the sequence features of PSAPLIPs and the evolution of PSAPLIP family in plants.

With the data currently available, only one or two *PSAPLIPs* genes occur in the genome in most species. However, some species may contain more copies, but all are less than 10 copies, such as rice and bananas. This may result from genome duplication during evolution or artificial selection in agriculture. And in terms of unique sequences, in most plant species, there are one to three *PSAPLIPs* genes across the genomes. This suggests that this family does not expand during evolution, but still persists in the genome, which indicates that this family is important in plant growth and development. The redundant alleles may disappear in natural selections.

Structural features of AtPSAPLIPs

Like PSAPLIPs in animals, plant PSAPLIPs also exhibit highly variable protein sequences, but the conservative cysteines remain the same and the distribution of cysteines and polar residues are well aligned with mammalian saposins. The high divergence can be visualized in the phylogenetic tree (Figure S15). Although most proteins are clustered into the major plant groups, many protein positions in the tree do not correspond to phylogenetic relationships between or among different plant groups. However, the aligned conservative cysteines and hydrophobic sites indicate that these SapB-like domains evolved from a single ancestor rather than independently. In most PSAPLIPs, all six cysteines can be found, and the distribution is conserved

(Figure 3-03). There are some sequences with only five conserved cysteines, but this is less likely to affect the overall structure. Cysteines are important for disulfide bonds, but they are not likely to determine the folding of the protein. The impact on mutation in conserved cysteines in different SAPLIPs is still unknown. As a result, the secondary structure can be predicted for these SapB-like domains.

With the sequence information, structures of *Arabidopsis* PSAPLIPs were constructed in Phyre2 and visualized in EzMol (Figure 3-01 and Figure 3-02). The overall secondary structure is highly similar with mammalian saposins and plant specific insert (PSI) in aspartic proteases. From the study of mammalian saposins, the primary function of these cysteines is forming disulfide bonds, which provide extra stability of protein configuration. This can be seen in the predicted structure of AtPSAPLIP1 (At3g51730) and AtPSAPLIP2 (At5g01800). Three pairs of conservative cysteines are shown in a direction advantaged for forming disulfide bonds (Figure 3-02). Overall, the highly divergent primary sequence and highly similar secondary structure is a feature among all saposin-like proteins in eukaryotes.

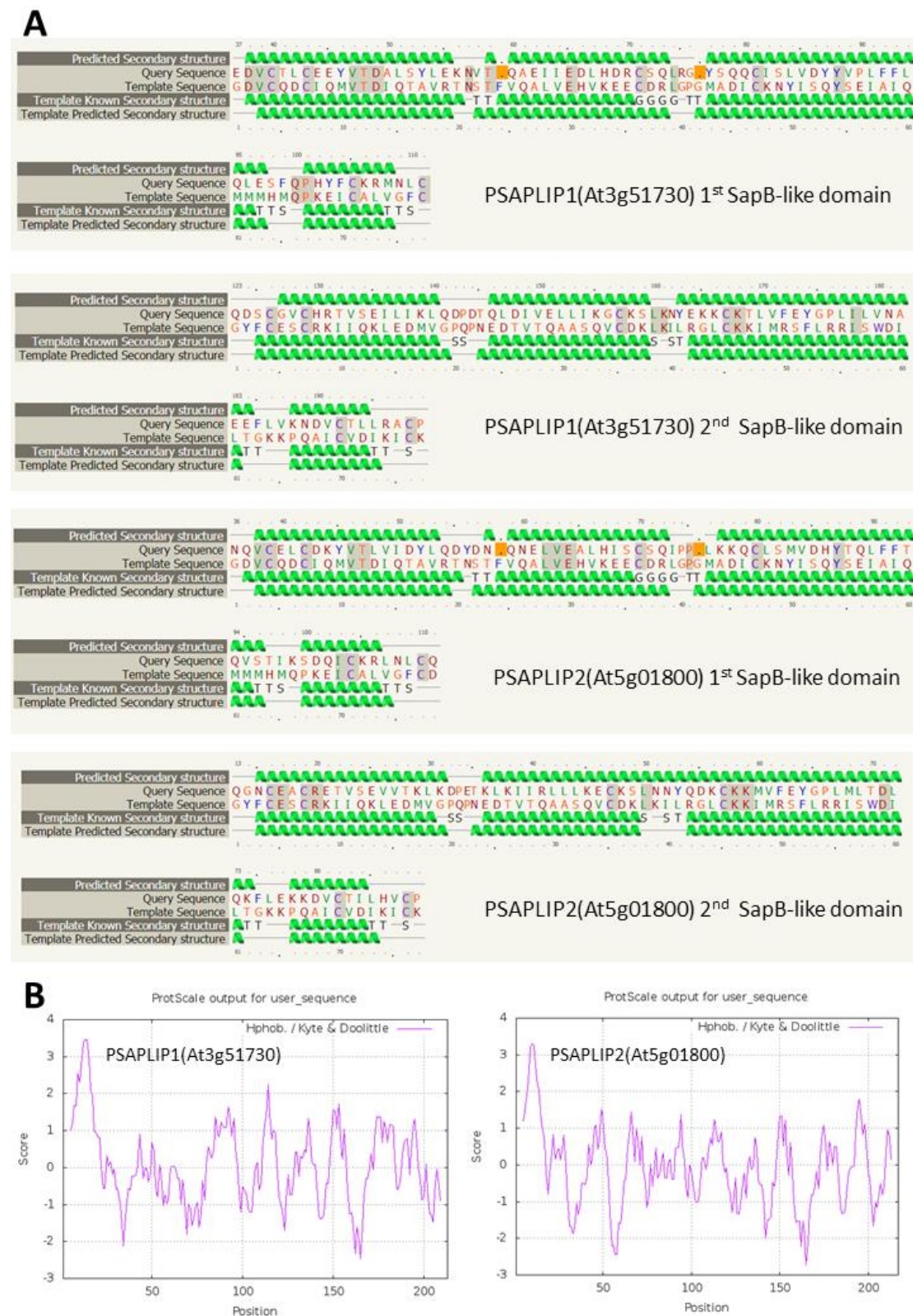


Figure 3-01. Predicted structure of AtPSAPLIP1 and AtPSAPLIP2. (A) Predicted secondary structure of AtPSAPLIP1 and AtPSAPLIP2. Alignment for the first saposin B (SapB)-like domain of AtPSAPLIP1 with chosen template is d1n69a. Confidence 99.66%.

Alignment for the second SapB-like domain of AtPSAPLIP1 with chosen template d1nkla. Confidence 99.71%. Alignment for the first SapB-like domain of AtPSAPLIP2 with chosen template is d1n69a. Confidence 99.79%. Alignment for the second SapB-like domain of AtPSAPLIP2 with chosen template d1nkla. Confidence 99.70%. Analysis was done by Phyre2. (B) Hydropathy plot for AtPSAPLIP1 and AtPSAPLIP2. Plots were created in ExPASy. Window size was 9 with linear weight variation model.

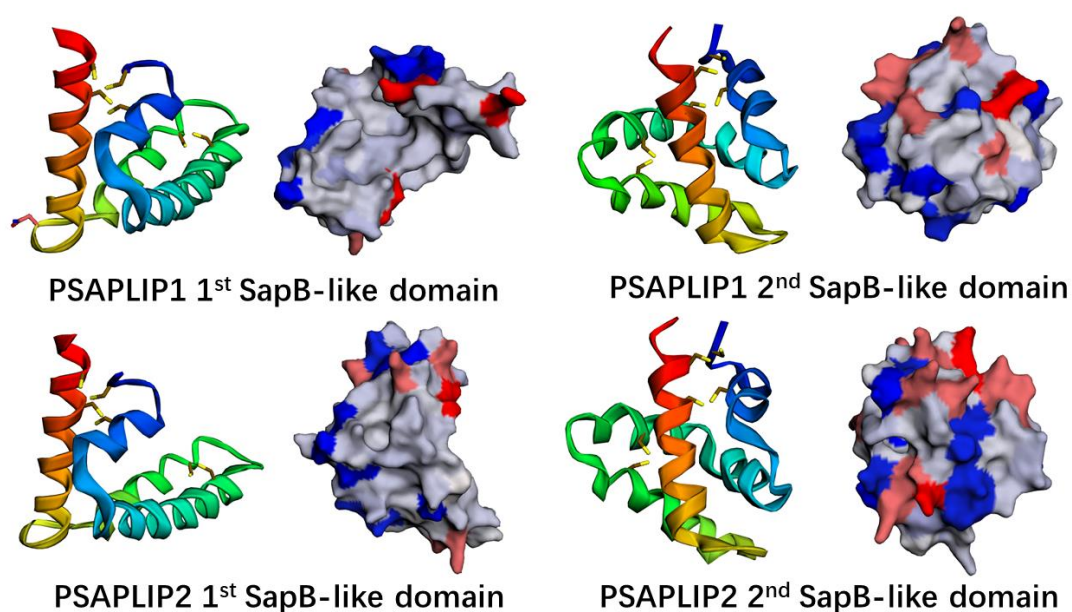


Figure 3-02. Predicted structure of saposin B (SapB)-like domains in AtPSAPLIP1 and AtPSAPLIP2. Image colored by rainbow N to C terminus with EzMol. Conserved cysteines are highlighted by brown sticks and the potential glycosylation site is highlighted by pink stick. In surface view, the negative region is colored by blue and positive region is colored by red.

The distribution of some of the conserved residues was slightly different between two SapB-like domains in the single PSAPLIP. For example, in the second SapB like

domain, the conserved aspartic site (D190 in AtPSAPLIP1) is close to the fifth cysteine (C192 in AtPSAPLIP1), while a conservative aspartic site (D86 in AtPSAPLIP1) is closer to the fourth cysteine (C81 in AtPSAPLIP1) and relative away from the fifth cysteine (C105 in AtPSAPLIP1) (Figure 3-02). The positively charged lysine (K155 in AtPSAPLIP1) is close to the third and fourth cysteines (C157 and C167 in AtPSAPLIP1) in the second SapB like domain but this does not present in the first one (Figure 3-03). As a result, conformation of these two SapB-like domains is likely to be slightly different. And it is also possible that two SapB-like domains are processed into two mature forms like human saposins. Based on mammalian saposins working model, saposins can form and function as dimers (Rossmann et al., 2008; Olmeda et al., 2012). In plant cells, the PSAPLIP might be processed into individual mature saposins like human prosaposin, or function on its own by bending and forming a 'self-dimer' with two lobes or forming a true dimer with another PSAPLIP molecule. Post-transcriptional processing into individual mature saposins is more likely in gymnosperms, liverworts, mosses and algae because there are three SapB-like domains in a single gene.

Unlike PSI in aspartic proteases, the direction of SapB-like domains in plant PSAPLIPs is not permuted. The PSI in aspartic proteases occur in green algae, which suggests that the saposin-like domains in aspartic proteases are very ancient and the relationship between PSI in modern aspartic proteases and modern PSAPLIPs is not close. PSI is likely to have evolved from an ancient duplication and deletion event. Therefore, aspartic proteases containing PSI and PSAPLIPs should be considered as two

different groups of proteins.

<i>Amborella_trichopoda</i> AMTR_s00007p00225690/1-214	
<i>Amborella_trichopoda</i> AMTR_s00062p00198130/1-320	
<i>Cinnamomum_micranthum_f._kanehirae</i> CKAN_01065200/1-278	
<i>Cinnamomum_micranthum_f._kanehirae</i> CKAN_00757300/1-212	
<i>Anthurium_amnicola</i> Psapl1_1/1-288	
<i>Anthurium_amnicola</i> Psapl1_2/1-273	
<i>Zostera_marina</i> ZOSMA_381G00120/1-242	
<i>Zostera_marina</i> ZOSMA_56G01350/1-232	
<i>Dendrobium_catenatum</i> MA16_Dca011512/1-222	
<i>Dendrobium_catenatum</i> MA16_Dca020165/1-215	
<i>Oryza_sativa_subsp._indica</i> Osl_34843/1-245	
<i>Oryza_sativa_subsp._indica</i> Osl_19500/1-223	
<i>Oryza_sativa_subsp._japonica</i> Osl2g0112200/1-245	
<i>Oryza_sativa_subsp._japonica</i> Osl5g0334400/1-223	
<i>Brachypodium_distachyon</i> BRADL_4g25580v3/1-245	
<i>Brachypodium_distachyon</i> BRADL_2g04110v3/1-235	
<i>Hordeum_vulgare_subsp._vulgare</i> NA F2DBE9/1-246	
<i>Hordeum_vulgare_subsp._vulgare</i> NA A0A287KA24/1-238	
<i>Triticum_aestivum</i> NA A0A3B6MMT5/1-246	
<i>Triticum_aestivum</i> NA A0A3B6FHG9/1-233	
<i>Sorghum_bicolor</i> SORBL_3008G032600/1-247	
<i>Sorghum_bicolor</i> SORBL_3003G055700/1-227	
<i>Zea_mays</i> ZEAMMB73_Zm00001d042734/1-240	
<i>Zea_mays</i> ZEAMMB73_Zm00001d039719/1-229	
<i>Aquilegia_coerulea</i> AQUUCO_00400489v1/1-223	
<i>Spinacia_oleracea</i> SOVF_050110/1-231	
<i>Helianthus_annuus</i> HannXRQ_Chrl0g028629/1-181	
<i>Cynara_cardunculus_var._scolymus</i> Cord_003008/1-231	
<i>Lactuca_sativa</i> LSAT_9X3806/1-229	
<i>Coffea_canephora</i> GSCOC_T0002323400/1-294	
<i>Nicotiana_tabacum</i> LOC107812754/1-245	
<i>Nicotiana_tabacum</i> LOC107792809/1-238	
<i>Solanum_tuberosum</i> 102602502/1-242	
<i>Solanum_lycopersicon</i> NA A0A3Q70I0/1-238	
<i>Cicer_arietinum</i> LOC101491522/1-279	1 MQLLNTLRWF 10
<i>Cicer_arietinum</i> LOC101508260/1-215	
<i>Medicago_truncatula</i> MTR_7g072560/1-242	
<i>Medicago_truncatula</i> MtrunA17_Chrl4g001314/1-223	
<i>Medicago_truncatula</i> MTR_029040/1-215	
<i>Trifoliumpratense</i> L195_g026334/1-194	
<i>Lotus_japonicus</i> NA 3S9R9/1-216	
<i>Phaseolus_vulgaris</i> PHAVU_008G084800g/1-222	
<i>Phaseolus_vulgaris</i> PHAVU_008G0847000g/1-217	
<i>Glycine_max</i> GLYMA_09G277100/1-237	
<i>Glycine_max</i> GLYMA_01G131400/1-216	
<i>Eucalyptus_grandis</i> EUGRSUZ_K01273/1-227	
<i>Eucalyptus_grandis</i> EUGRSUZ_A00687/1-219	
<i>Gossypium_hirsutum</i> LOC107896756/1-233	
<i>Gossypium_hirsutum</i> LOC107935966/1-227	
<i>Gossypium_tomentosum</i> ES332_D10G139500v1/1-233	
<i>Gossypium_tomentosum</i> ES332_A02G005200v1/1-227	
<i>Theobroma_cacao</i> TCM_019744/1-228	
<i>Brassica_rapa_subsp._pekinensis</i> NA M4D8NC/1-215	
<i>Brassica_rapa_subsp._pekinensis</i> NA M4CRM9/1-214	
<i>Brassica_oleracea_var._oleracea</i> NA A0A0D3DSC3/1-229	
<i>Brassica_oleracea_var._oleracea</i> NA A0A0D3D313/1-216	
<i>Arabidopsis_lyrata_subsp._lyrata</i> ARALYDRAFT_486888/1-220	
<i>Arabidopsis_lyrata_subsp._lyrata</i> ARALYDRAFT_666001/1-213	
<i>Arabidopsis_thaliana</i> At5g01800/1-217	
<i>Arabidopsis_thaliana</i> At3g51730/1-213	
<i>Rosa_chinensis</i> RchiOBHm_Chrl3g046096/1-229	
<i>Prunus_persica</i> PRUPE_6G290000/1-253	
<i>Malus_domestica</i> DVH24_036312/1-296	1 MGNWVAGC I KQRSKANDHSRPFNAPPNAQPQQRKKGG IP 39	
<i>Populus_trichocarpa</i> POPTR_016G133400/1-242	
<i>Populus_trichocarpa</i> POPTR_006G107300/1-242	
<i>Cucumis_sativus</i> Csa_4G331080/1-233	
<i>Cucumis_melo_var._makuwa</i> E5676_scaffold127G001120/1-249	
<i>Cucumis_melo_var._makuwa</i> E6C27_scaffold1166G00310/1-233	

Conservation



Quality



Consensus

MGNWVAGC I KQRSKANDHSRPFNAPPNAQ+Q+++++++

Occupancy



Amborella_trichopoda AMTR_s00007p00225690/1-214	1	MGVGRTRF	7
Amborella_trichopoda AMTR_s00062p00198130/1-320	1	..MVSLERSVRHPQSQSLFFPVLGWVF	TVKMRCT	32
Cinnamomum_micranthum_f_kanehirae CKAN_01065200/1-278	1	MGM	3
Cinnamomum_micranthum_f_kanehirae CKAN_00757300/1-212	1	MSVSI RL	7
Anthurium_amicola Psapl1_1/1-288	1	MKMGLVPSL	9
Anthurium_amicola Psapl1_2/1-273	1	MMDNITM	7
Zostera_marina ZOSMA_381G00120/1-242	1	MLRMGLQLFLV	11
Zostera_marina ZOSMA_56G01350/1-232	1	MDGRLSL	7
Dendrobium_catenatum MA16_Dca011512/1-222	1	MDLKEG	6
Dendrobium_catenatum MA16_Dca020165/1-215	1	MGSKAAL	7
Oryza_sativa_subsp_indica Osl_34843/1-245	1	MDFKVA	6
Oryza_sativa_subsp_indica Osl_19500/1-223	1	MGPRVVF	7
Oryza_sativa_subsp_japonica Osl2g0112200/1-245	1	MDFKVAS	7
Oryza_sativa_subsp_japonica Osl2g05g0334400/1-223	1	MGPRVVF	7
Brachypodium_distachyon BRADL_4g25580v3/1-245	1	MGLGLRVPV	9
Brachypodium_distachyon BRADL_2g04110v3/1-235	1	MACLTRR	7
Hordeum_vulgare_subsp_vulgare NA F2DBE9/1-246	1	MGLGLGAPF	9
Hordeum_vulgare_subsp_vulgare NA A0A287KA24/1-238	1	MAMACSTRM	9
Triticum_aestivum NA A0A3B6MMT5/1-246	1	MGLGLGAP	8
Triticum_aestivum NA A0A3B6FHG9/1-233	1	MASSTRR	7
Sorghum_bicolor SORBL_3008G032600/1-247	1	MGSKAPLC	8
Sorghum_bicolor SORBL_3003G055700/1-227	1	MCSITRL	7
Zea_mays ZEAAMMB73_Zm00001d042734/1-240	1	MGSKAPF	7
Zea_mays ZEAAMMB73_Zm00001d039719/1-229	1	MCSVARL	7
Aquilegia_coerulea AQUUCO_00400489v1/1-223	1	MGVKVGL	7
Spinacia_oleracea SOV_050110/1-231	1	MDVRVGL	7
Helianthus_annuus HannXRQ_Chr10g028629/1-181	1	MGKK	4
Cynara_cardunculus_var_scolymus Cord_003008/1-231	1	MGRIGL	7
Lactuca_sativa LSAT_9X3806/1-229	1	MGKGLGL	7
Coffea_canephora GSCOC_70002323400/1-294	1MCLSLLLFFKIVKSNHQSI	P LELKLAKRPFPS	31
Nicotiana_tabacum LOC107812754/1-245	1	MDFRVC	6
Nicotiana_tabacum LOC107792809/1-238	1	MDVKVC	6
Solanum_tuberosum 102602502/1-242	1	MDLRLC	6
Solanum_lycopersicon NA A0A3Q700/1-238	1	MDLRLC	6
Cicer_arietinum LOC101491522/1-279	11	CSFKYKRKNVSSYIVYDSDLLKFVIRSM	GGRIGL	44
Cicer_arietinum LOC101508260/1-215	1	MEGKIGL	7
Medicago_truncatula MTTR_7g072560/1-242	1	MEGRIGL	7
Medicago_truncatula MTtrunA17_Chr4g001314/1-223	1MYIFTLERN	S	10
Medicago_truncatula MTTR_029040/1-215	1	MEGKIGF	7
Trifoliumpratense L195_g026334/1-194	1MQQ	3
Lotus_japonicus NA J3S9R9/1-216	1	MKGRMGL	7
Phaseolus_vulgaris PHAVU_008G084800g/1-222	1	MEGRMGL	7
Phaseolus_vulgaris PHAVU_008G0847000g/1-217	1	MEGRMTL	7
Glycine_max GLYMA_09G277100/1-237	1	MEGRMGL	7
Glycine_max GLYMA_01G131400/1-216	1	MEERVGI	7
Eucalyptus_grandis EUGRSUZ_K01273/1-227	1	MGVKIGF	7
Eucalyptus_grandis EUGRSUZ_A00687/1-219	1	MEGKIGL	7
Gossypium_hirsutum LOC107896756/1-233	1	MLSKGIMDVRVGL	13
Gossypium_hirsutum LOC107935966/1-227	1	MDARFGL	7
Gossypium_tomentosum ES332_D10G139500v1/1-233	1	MLSKGIMDVRVGL	13
Gossypium_tomentosum ES332_A02G005200v1/1-227	1	MDARFGL	7
Theobroma_cacao TCM_019744/1-228	1	MDARVGL	7
Brassica_rapa_subsp_pekinensis NA M4D8N0/1-215	1	MGPKAGT	7
Brassica_rapa_subsp_pekinensis NA M4CRM9/1-214	1	MGPKAGT	7
Brassica_oleracea_var_oleracea NA A0A0D3DSC3/1-229	1	MGPKAGT	7
Brassica_oleracea_var_oleracea NA A0A0D3D313/1-216	1	MGPKAGT	7
Arabidopsis_lyrata_subsp_lyrata ARALYDRAFT_486888/1-220	1	MGGRFGV	7
Arabidopsis_lyrata_subsp_lyrata ARALYDRAFT_66600/1-213	1	MGLKAGT	7
Arabidopsis_thaliana At5g01800/1-217	1	MGGRFGV	7
Arabidopsis_thaliana At3g51730/1-213	1	MGLKAGT	7
Rosa_chinensis RchiOBHm_Chr3g046096/1-229	1	MDMRVGF	7
Prunus_persica PRUPE_6G290000/1-253	1	MDVRVGV	7
Malus_domestica DVH24_036312/1-296	40	RHQVFHFFPSLSQSQVNIYLSISDCT	MDMRVGV	73
Populus_trichocarpa POPTR_016G133400/1-242	1	MDLRIGL	7
Populus_trichocarpa POPTR_006G107300/1-242	1	MDMRIGL	7
Cucumis_sativus Csa_4G331080/1-233	1	MDLRF AI	7
Cucumis_melo_var_makuwa E5676_scaffold127G001120/1-249	1	MKLKL-W	6
Cucumis_melo_var_makuwa E6C27_scaffold1166G00310/1-233	1	MDSRF AI	7

Conservation



Quality



Consensus



Occupancy



<i>Amborella trichopoda</i> AMTR_s00007p00225690/1-214	8LFLLL..LLTI...	15
<i>Amborella trichopoda</i> AMTR_s00062p00198130/1-320	33LPLVVTFLLII...	42
<i>Cinnamomum micranthum</i> _f_kanehirae CKAN_01065200/1-278	4LSFFLGILVL...	13
<i>Cinnamomum micranthum</i> _f_kanehirae CKAN_00757300/1-212	8LFLLL..LLSS...	15
<i>Anthurium amnicola</i> Psapl1_1/1-288	10AVLL..FLAV...	17
<i>Anthurium amnicola</i> Psapl1_2/1-273	8LMYVVMMLSI...	17
<i>Zostera marina</i> ZOSMA_381G00120/1-242	12VLFFIVNLGT...	21
<i>Zostera marina</i> ZOSMA_56G01350/1-232	8LFLI...ILGI...	15
<i>Dendrobium catenatum</i> MA16_Dca011512/1-222	7IFLIVLLIGS...	16
<i>Dendrobium catenatum</i> MA16_Dca020165/1-215	8IFLG...IMI...	15
<i>Oryza sativa</i> _subsp_indica Osl_34843/1-245	7SFLLLLLIVT...	16
<i>Oryza sativa</i> _subsp_indica Osl_19500/1-223	8MFIIAVMLLG...	17
<i>Oryza sativa</i> _subsp_japonica Os12g0112200/1-245	8FLLL...LLIV...	15
<i>Oryza sativa</i> _subsp_japonica Os05g0334400/1-223	8MFIIAVMLLG...	17
<i>Brachypodium distachyon</i> BRADL_4g25580v3/1-245	10SFLLLLLLV...	19
<i>Brachypodium distachyon</i> BRADL_2g04110v3/1-235	8LTFLLVVLAFS...	17
<i>Hordeum vulgare</i> _subsp_vulgare NA F2DBE9/1-246	10FFFLLILLAA...	19
<i>Hordeum vulgare</i> _subsp_vulgare NA A0A287KA24/1-238	10LAFLLALAF...	19
<i>Triticum aestivum</i> NA A0A3B6MMT5/1-246	9FLLVLVLLLA...	18
<i>Triticum aestivum</i> NA A0A3B6FHG9/1-233	8LAFVLLVLAF...	17
<i>Sorghum bicolor</i> SORBL_3008G032600/1-247	9LTLLLLLLVA...	18
<i>Sorghum bicolor</i> SORBL_3003G055700/1-227	8AFVLALAIAS...	17
<i>Zea mays</i> ZEMMB73_Zm00001d042734/1-240	8CLTLLLLLAA...	17
<i>Zea mays</i> ZEMMB73_Zm00001d039719/1-229	8AFVLALAIAS...	17
<i>Aquilegia coerulea</i> AQUCO_00400489v1/1-223	8LFLLL..VLGI...	15
<i>Spinacia oleracea</i> SOVF_050110/1-231	8VVLL..VVG...	15
<i>Helianthus annuus</i> HannXRQ_Ch10g0286291/1-181		
<i>Cynara cardunculus</i> _var_scolymus Cord_003008/1-231	8IFVF...LLVA...	15
<i>Lactuca sativa</i> LSAT_9X38061/1-229	8VFVF...LLAV...	15
<i>Coffea canephora</i> GSCOC_T00023234001/1-294	32	LVFPPPLQLYCIHQIFCKNSEMDVWVPFLFLL..LLSN...	65
<i>Nicotiana tabacum</i> LOC107812754/1-245	7LCLF...ILGS...	14
<i>Nicotiana tabacum</i> LOC107792809/1-238	7LCLL...ILGA...	14
<i>Solanum tuberosum</i> 102602502/1-242	7LCLF...ILGS...	14
<i>Solanum lycopersicum</i> NA A0A3Q7000/1-238	7LCLF...IIGS...	14
<i>Cicer arietinum</i> LOC101491522/1-279	45LFLV..VVG...	52
<i>Cicer arietinum</i> LOC101508260/1-215	8LFMI...VLVS...	15
<i>Medicago truncatula</i> MTR_7g072560/1-242	8LFIV...VLGF...	15
<i>Medicago truncatula</i> MtrunA17_Ch14g001314/1-223	11LCLN...IVSV...	18
<i>Medicago truncatula</i> MTR_029040/1-215	8LFLI...LLGA...	15
<i>Trifolium pratense</i> L195_g026334/1-194	4LYLK...LYG...	10
<i>Lotus japonicus</i> NA 3S9R9/1-216	8LFLI...VLGT...	15
<i>Phaseolus vulgaris</i> PHAVU_008G084800g/1-222	8LFLV...VLGA...	15
<i>Phaseolus vulgaris</i> PHAVU_008G0847000g/1-217	8LFLV...VLGA...	15
<i>Glycine max</i> GLYMA_09G277100/1-237	8LFLV...VLGA...	15
<i>Glycine max</i> GLYMA_01G131400/1-216	8LFLV...VLGA...	15
<i>Eucalyptus grandis</i> EUGRSUZ_K01273/1-227	8LFLV...VLGS...	15
<i>Eucalyptus grandis</i> EUGRSUZ_A00687/1-219	8LFLV...LWST...	15
<i>Gossypium hirsutum</i> LOC107896756/1-233	14LFLF...VLGA...	21
<i>Gossypium hirsutum</i> LOC107935966/1-227	8LFLF...MLGA...	15
<i>Gossypium tomentosum</i> ES332_D10G139500v1/1-233	14LFLF...VLGA...	21
<i>Gossypium tomentosum</i> ES332_A02G005200v1/1-227	8LFLF...MLGA...	15
<i>Theobroma cacao</i> TCM_019744/1-228	8LFLF...VLGA...	15
<i>Brassica rapa</i> _subsp_pekinensis NA M4D8N0/1-215	8FVLF...LLGL...	15
<i>Brassica rapa</i> _subsp_pekinensis NA M4CRM9/1-214	8LVII...LLGL...	15
<i>Brassica oleracea</i> _var_oleracea NA A0A0D3DSC3/1-229	8LVII...LLGL...	15
<i>Brassica oleracea</i> _var_oleracea NA A0A0D3D313/1-216	8FVLF...LLGL...	15
<i>Arabidopsis lyrata</i> _subsp_lyrata ARALYDRAFT_486888/1-220	8LFLFLGLLTC...	17
<i>Arabidopsis lyrata</i> _subsp_lyrata ARALYDRAFT_666001/1-213	8FVLF...LLGL...	15
<i>Arabidopsis thaliana</i> At5g01800/1-217	8LLVL...FLLS...	15
<i>Arabidopsis thaliana</i> At3g51730/1-213	8FVLL...LLGL...	15
<i>Rosa chinensis</i> RchiOBHm_Ch3g0460961/1-229	8ILFV...VLGA...	15
<i>Prunus persica</i> PRUPE_6G290000/1-253	8LVLF...VLGA...	15
<i>Malus domestica</i> DVH24_036312/1-296	74LFLF...VLAA...	81
<i>Populus trichocarpa</i> POPTR_016G133400/1-242	8LFLV...ALGA...	15
<i>Populus trichocarpa</i> POPTR_006G107300/1-242	8LFLV...TLA...	15
<i>Cucumis sativus</i> Csa_4G331080/1-233	8VFLV...VLGV...	15
<i>Cucumis melo</i> _var_makuwa E5676_scaffold127G001120/1-249	7LWLL...VLGLPPF...	15
<i>Cucumis melo</i> _var_makuwa E6C27_scaffold1166G00310/1-233	8VFLV...VLVS...	15

Conservation



Quality



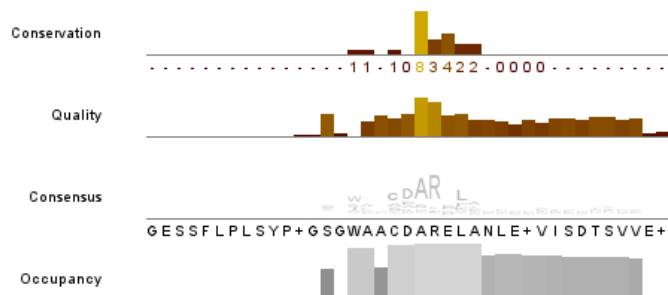
Consensus

LFLV...VLG
LVFPPLQLYCIHQIFCKNSEMDVWVPLFLLLLLLLGAAPPF

Occupancy



Amborella_trichopoda|AMTR_s00007p00225690/1-214
Amborella_trichopoda|AMTR_s00062p00198130/1-320
Cinnamomum_micranthum_f_kanehirae|CKAN_01065200/1-278
Cinnamomum_micranthum_f_kanehirae|CKAN_00757300/1-212
Anthurium_amnicola|Psapl_1_1-288
Anthurium_amnicola|Psapl_1_2/1-273
Zostera_marina|ZOSMA_381G00120/1-242
Zostera_marina|ZOSMA_56G01350/1-232
Dendrobium_catenatum|MA16_Dca011512/1-222
Dendrobium_catenatum|MA16_Dca020165/1-215
Oryza_sativa_subsp_indica|Osl_34843/1-245
Oryza_sativa_subsp_indica|Osl_19500/1-223
Oryza_sativa_subsp_japonica|Osl_0112200/1-245
Oryza_sativa_subsp_japonica|Osl_05g0334400/1-223
Brachypodium_distachyon|BRADI_4g25580v3/1-245
Brachypodium_distachyon|BRADI_2g04110v3/1-235
Hordeum_vulgare_subsp_vulgare|NA|F2DBE9/1-246
Hordeum_vulgare_subsp_vulgare|NA|A0A287KA24/1-238
Triticum_aestivum|NA|A0A3B6MMT5/1-246
Triticum_aestivum|NA|A0A3B6FHG9/1-233
Sorghum_bicolor|SORBL_3008G032600/1-247
Sorghum_bicolor|SORBL_3003G055700/1-227
Zea_mays|ZEAMMB73_Zm00001d042734/1-240
Zea_mays|ZEAMMB73_Zm00001d039719/1-229
Aquilegia_coerulea|AQUUCO_00400489v1/1-223
Spinacia_oleracea|SOVF_050110/1-231
Helianthus_annuus|HannXRG_Chr10g028629/1-181
Cynara_cardunculus_var_scolymus|Ccd_003008/1-231
Lactuca_sativa|LSAT_9X3806/1-229
Coffea_canephora|GSCOC_T0002323400/1-294
Nicotiana_tabacum|LOC107812754/1-245
Nicotiana_tabacum|LOC107792809/1-238
Solanum_tuberosum|102602502/1-242
Solanum_lycopersicon|NA|A0A3Q7010/1-238
Cicer_arietinum|LOC101491522/1-279
Cicer_arietinum|LOC101508260/1-215
Medicago_truncatula|MTR_7g072560/1-242
Medicago_truncatula|MtrunA17_Chr4g001314/1-223
Medicago_truncatula|MTR_029040/1-215
Trifoliumpratense|L195_g026334/1-194
Lotus_japonicus|NA|3S9R9/1-216
Phaseolus_vulgaris|PHAVU_008G084800g/1-222
Phaseolus_vulgaris|PHAVU_008G0847000g/1-217
Glycine_max|GLYMA_09G277100/1-237
Glycine_max|GLYMA_01G131400/1-216
Eucalyptus_grandis|EUGRSUZ_K01273/1-227
Eucalyptus_grandis|EUGRSUZ_A00687/1-219
Gossypium_hirsutum|LOC107896756/1-233
Gossypium_hirsutum|LOC107935966/1-227
Gossypium_tomentosum|ES332_D10G139500v1/1-233
Gossypium_tomentosum|ES332_A02G005200v1/1-227
Theobroma_cacao|TCM_019744/1-228
Brassica_rapa_subsp_pekinensis|NA|M4D8NQ/1-215
Brassica_rapa_subsp_pekinensis|NA|M4CRM9/1-214
Brassica_oleracea_var_oleracea|NA|A0A0D3DSC3/1-229
Brassica_oleracea_var_oleracea|NA|A0A0D3D313/1-216
Arabidopsis_lyrata_subsp_lyrata|ARALYDRAFT_486888/1-220
Arabidopsis_lyrata_subsp_lyrata|ARALYDRAFT_666001/1-213
Arabidopsis_thaliana|At5g01800/1-217
Arabidopsis_thaliana|At3g51730/1-213
Rosa_chinensis|RchiOBHm_Chr3g046096/1-229
Prunus_persica|PRUPE_6G290000/1-253
Malus_domestica|DVH24_036312/1-296
Populus_trichocarpa|POPTR_016G133400/1-242
Populus_trichocarpa|POPTR_006G107300/1-242
Cucumis_sativus|Csa_4G331080/1-233
Cucumis_melo_var_makuwa|E5676_scaffold127G001120/1-249
Cucumis_melo_var_makuwa|E6C27_scaffold1166G00310/1-233



<i>Amborella_trichopoda</i> AMTR_s00007p00225690/1-214	39TFTLITR---D--ERV--	CTYCE	54
<i>Amborella_trichopoda</i> AMTR_s00062p00198130/1-320	67	RSFDGFIEGSSSRIEIPSLKEF---PL-EFI--	CNACL	99
<i>Cinnamomum_micranthum_f_kanehirae</i> CKAN_01065200/1-278	36LAGPRTL DGLPP-.....EFF--	CNYCL	55
<i>Cinnamomum_micranthum_f_kanehirae</i> CKAN_00757300/1-212	38EDPRSKTLEVDVG---N--ERL--	CTYCE	59
<i>Anthurium_amnicola</i> Psapl_1/1-288	42	DSLVLHELKEPYENEVLSSSEGLP-.....TEF--	CHLCL	72
<i>Anthurium_amnicola</i> Psapl_1/1-273	41KKDNHSLLEVIMT-.....NYF--	CNSCI	60
<i>Zostera_marina</i> ZOSMA_381G00120/1-242	44NFIPNTKLSDESKK---S--GFF--	CDTCL	66
<i>Zostera_marina</i> ZOSMA_56G01350/1-232	39SEMPITEIKPAKFNS---N--KDTVY	CTLCE	64
<i>Dendrobium_catenatum</i> MA16_Dca011512/1-222	27GENPITEGKGGE---P--EQL--	CTVCE	47
<i>Dendrobium_catenatum</i> MA16_Dca020165/1-215	27FMDPKVTAT---N--GQI--	CQNCL	44
<i>Oryza_sativa_subsp_indica</i> Osl_34843/1-245	38	GEDASPMYKEQIALT KIPVTL LR--SKHSSL--	CSACE	71
<i>Oryza_sativa_subsp_indica</i> Osl_19500/1-223	35ETSLTMKV---D--PQL--	CQICE	51
<i>Oryza_sativa_subsp_japonica</i> Osl_204110v3/1-235	38	GEDASPMYKEQIALT KIPVTL LR--SKHSSL--	CSACE	71
<i>Oryza_sativa_subsp_japonica</i> Osl_05g0334400/1-223	35ETSLTMKV---D--PQL--	CQICE	51
<i>Brachypodium_distachyon</i> BRADL_4g25580v3/1-245	42GQTYKEQIISSKIPVHVER---G--NPL--	CSACK	69
<i>Brachypodium_distachyon</i> BRADL_2g04110v3/1-235	40RESLPLVSKGAGLTAA---N--GKL--	CVLCE	64
<i>Hordeum_vulgare_subsp_vulgare</i> NA F2DBE9/1-246	45HPNDKEEITSSKIPVSVES---G--TTV--	CTSTCE	72
<i>Hordeum_vulgare_subsp_vulgare</i> NA A0A287KA24/1-238	43QESFPLASKGAGLTSA---N--GKL--	CVLCE	67
<i>Triticum_aestivum</i> NA A0A386MMT5/1-246	45RPNGKEEISSKIHVSVES---G--STI--	CTSTCE	72
<i>Triticum_aestivum</i> NA A0A386FHG9/1-233	41QESFPLASRAGLTSA---N--GKL--	CVLCE	65
<i>Sorghum_bicolor</i> SORBL_3008G032600/1-247	43	PDHGSTLKEQISSMKTPVHLKS---S--GQI--	CLACE	73
<i>Sorghum_bicolor</i> SORBL_3003G055700/1-227	41GPGLTAA---S--GKL--	CQLCE	56
<i>Zea_mays</i> ZEAMMB73_Zm00001d042734/1-240	38LKEHISSTKIPARLKR---G--SGL--	CSACE	62
<i>Zea_mays</i> ZEAMMB73_Zm00001d039719/1-229	42GSSGLAAT---S--GKL--	CQLCE	58
<i>Aquilegia_coerulea</i> AQUCO_00400489v1/1-223	38QKGSK---D--DKV--	CTMCE	51
<i>Spinacia_oleracea</i> SOVFL_050110/1-231	38LRGPGHEHRG---P--IDV--	CTMCE	55
<i>Helianthus_annuus</i> HannXRQ_Chr10g028629/1-181	5DEL--	CSLCE	12
<i>Cynara_cardunculus_var_scolymus</i> Cord_003008/1-231	38LRSKAEKRFGLGNVKN---DNL--	CSLCE	62
<i>Lactuca_sativa</i> LSATL_9X3806/1-229	38LQSKAAKRFGLGNVKN---DNL--	CTLCE	63
<i>Coffea_canephora</i> GSCOC_T0002323400/1-294	87	ASDFQLNGQQPNKEVQTADGFDQ---N--DQV--	CMLCE	118
<i>Nicotiana_tabacum</i> LOC107812754/1-245	37	SVLQINNLEVP RQVPPEEVGG---N--EQL--	CTLCE	67
<i>Nicotiana_tabacum</i> LOC107792809/1-238	36SALWISNLQAQKQLQSLKDV---EGL--	CTLCE	63
<i>Solanum_tuberosum</i> 102602502/1-242	37	SVLQINNLEELRQVQPLEEVNG---N--EQL--	CTLCE	67
<i>Solanum_lycopersicon</i> NA A0A3Q700/1-238	37QINNLEELRQVQPLEEVNG---S--EQL--	CTLCE	64
<i>Cicer_arietinum</i> LOC101491522/1-279	75ASSELGR---I--PDV--	CALCE	90
<i>Cicer_arietinum</i> LOC101508260/1-215	26NPELNI---T--SDV--	CSLCE	40
<i>Medicago_truncatula</i> MTR_7g072560/1-242	38SELGR---I--PDV--	CALCE	51
<i>Medicago_truncatula</i> MtrunA17_Chr4g001314/1-223	33SYAELNR---K--PDA--	CSICE	48
<i>Medicago_truncatula</i> MTR_029040/1-215	26NPELNR---K--PDA--	CSICE	40
<i>Trifoliumpratense</i> L195_g026334/1-194	11K--SDA--	CTICE	19
<i>Lotus_japonicus</i> NA J3S9R9/1-216	26NPELNR---K--SDV--	CALCE	40
<i>Phaseolus_vulgaris</i> PHAVL_008G084800g/1-222	38GISELKT---K--LDM--	CALCE	53
<i>Phaseolus_vulgaris</i> PHAVL_008G084700g/1-217	26NRDHF I K L I R---K--PDA--	CALCE	44
<i>Glycine_max</i> GLYMA_09G277100/1-237	26KPDLLSKLSR---K--PDA--	CALCE	44
<i>Glycine_max</i> GLYMA_01G131400/1-216	26IISELNR---K--SDV--	CELCE	41
<i>Eucalyptus_grandis</i> EUGRSUZ_K01273/1-227	38KGSST---N--GYM--	CEWCE	51
<i>Eucalyptus_grandis</i> EUGRSUZ_A00687/1-219	26HSIVRGEGLR---N--DNV--	CMLCE	44
<i>Gossypium_hirsutum</i> LOC107896756/1-233	44QTNQGQDEEVVENI VW---K--DNV--	CTLCE	68
<i>Gossypium_hirsutum</i> LOC107935966/1-227	38QVNWRQDEKVIETVAR---N--DNV--	CTLCE	62
<i>Gossypium_tomentosum</i> ES332_D10G139500v1/1-233	44QTNQGQDEEVVENI VG---K--DNV--	CTLCE	68
<i>Gossypium_tomentosum</i> ES332_A02G005200v1/1-227	38QVNWRQDEKVIETVAR---N--DNV--	CTLCE	62
<i>Theobroma_cacao</i> TCM_019744/1-228	38QINQGQDEEVVKKVAR---N--DNV--	CTLCE	62
<i>Brassica_rapa_subsp_pekinensis</i> NA M4DBN0/1-215	26HPTLSEEVTK---N--ENV--	CTLCE	44
<i>Brassica_rapa_subsp_pekinensis</i> NA M4CRM9/1-214	31SEKVIK---N--EKV--	CTLCE	45
<i>Brassica_oleracea_var_oleracea</i> NA A0A0D3DSC3/1-229	38NLFLFD FCAEKVIK---N--EKV--	CTLCE	60
<i>Brassica_oleracea_var_oleracea</i> NA A0A0D3D313/1-216	26HQSTLSEEVSK---N--ENV--	CTLCE	45
<i>Arabidopsis_lyrata_subsp_lyrata</i> ARALYDRAFT_486888/1-220	39NQV--	CELCD	46
<i>Arabidopsis_lyrata_subsp_lyrata</i> ARALYDRAFT_666001/1-213	26DSTLSEKVS N---K--EDV--	CTLCE	44
<i>Arabidopsis_thaliana</i> At5g01800/1-217	36NQV--	CELCD	43
<i>Arabidopsis_thaliana</i> At3g51730/1-213	26DSTI SEKVS N---K--EDV--	CTLCE	44
<i>Rosa_chinensis</i> RchiOBHm_Chr3g0460961/1-229	32QEGEPQTLKEFSG---N--ENV--	CTLCE	53
<i>Prunus_persica</i> PRUPE_6G290000/1-253	38IQVVE TQT LQVSGDEFVG---N--DNV--	CTLCE	64
<i>Malus_domestica</i> DVH24_036312/1-296	104WETQTFQVEEVVG---N--DNV--	CTLCE	125
<i>Populus_trichocarpa</i> POPTR_016G133400/1-242	39	ISVEIMENQEQENEIQT TNNVTR---K--DEV--	CTLCE	70
<i>Populus_trichocarpa</i> POPTR_006G107300/1-242	39	ISA I KMKNQEQETDIQT SNNVTR---K--DEV--	CTLCE	70
<i>Cucumis_sativus</i> Csa_4G331080/1-233	38KDVEALSEASS---N--SKI--	CTLCE	57
<i>Cucumis_melo_var_makuwa</i> E5676_scaffold127G001120/1-249	53EKDVEALSEASS---N--PKI--	CTLCE	73
<i>Cucumis_melo_var_makuwa</i> E6C27_scaffold1166G00310/1-233	38KDVEALSEASS---N--PKI--	CTLCE	57

Conservation



Quality



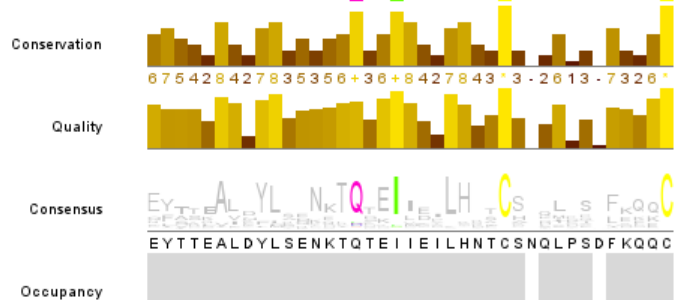
Consensus



Occupancy

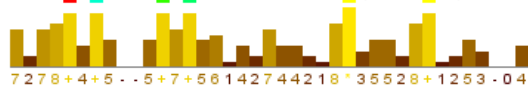


<i>Amborella_trichopoda</i> AMTR_s00007p00225690/1-214	55 QF ASEAF EYLGNNQTQTD I I KTLHQVCS - SMYS - FKHQC	91
<i>Amborella_trichopoda</i> AMTR_s00062p00198130/1-320	100 EALRLAEKVLADPEFL EN I KKCAGDI CS - LLPSNLQGE C	137
<i>Cinnamomum_micranthum_f_kanehirae</i> CKAN_01065200/1-278	56 S I SRDVEKVLADPKLLEKASMIASELCH - ILPSDLQKKC	93
<i>Cinnamomum_micranthum_f_kanehirae</i> CKAN_00757300/1-212	60 QF TAQAIYYLSENKTQSEIVEALHHTCS - RLRT - FHKEC	96
<i>Anthurium_amnicola</i> Psapl_1_1/1-288	73 DF SRKAKKILSDPNLKEIDNLAAVLCS - LVSPDLKPKC	110
<i>Anthurium_amnicola</i> Psapl_1_2/1-273	61 EASAKVEEILIDESFNEEIDALSNEICHNRVPSIEMKMC	99
<i>Zostera_marina</i> ZOSMA_381G00120/1-242	67 DVSSEVTKILSDSDLSDKIEVFFAQLCQ - ILPSDMESKC	104
<i>Zostera_marina</i> ZOSMA_56G01350/1-232	65 EYASLADYLSQNKQTQTEI IDSLTQACS - RLKS - FQQQC	101
<i>Dendrobium_catenatum</i> MA16_Dca011512/1-222	48 VFTARATTFLENENKTSEILDTLHHACS - ELRS - LELKC	84
<i>Dendrobium_catenatum</i> MA16_Dca020165/1-215	45 EFTTKAILFLGKNETQTEIIGNLHQACS - HLLS - FEPQC	81
<i>Oryza_sativa_subsp_indica</i> Osl_34843/1-245	72 NITSEAVNFLSEKQIQDKIMTILHDTCS - QTFS - FEQKC	108
<i>Oryza_sativa_subsp_indica</i> Osl_19500/1-223	52 EFATEALFYLNENETQVEI IATLHQACS - KFPS - FKLEC	88
<i>Oryza_sativa_subsp_japonica</i> Osl_0112200/1-245	72 NITSEAVNFLSEKQIQDKIMTILHDTCS - QTFS - FEQKC	108
<i>Oryza_sativa_subsp_japonica</i> Osl_05g0334400/1-223	52 EFATEALFYLNENETQVEI IATLHQACS - KFPS - FKLEC	88
<i>Brachypodium_distachyon</i> BRADL_4g25580v3/1-245	70 NLTNEAVSYLSQKQSQDKMLEVLHEACS - QTFS - LEQKC	106
<i>Brachypodium_distachyon</i> BRADL_2g04110v3/1-235	65 QYSTEALFYLQNKETQTEILSVLHHGCA - NLGP - LRQQC	101
<i>Hordeum_vulgare_subsp_vulgare</i> NA F2DBE9/1-246	73 NLTNKSVSYSLEKQQTDEIMEILHGACS - QTFS - LEQKC	109
<i>Hordeum_vulgare_subsp_vulgare</i> NA A0A287KA24/1-238	68 QYSTEALVYLRQKETQTEILSVLHHTCA - SLGP - LRQQC	104
<i>Triticum_aestivum</i> NA A0A3B6MMT5/1-246	73 NLTNKAVSYSLEKQQTDEIMEILHGACS - QTFS - LEQKC	109
<i>Triticum_aestivum</i> NA A0A3B6FHG9/1-233	66 QYSTEALVYLRQKETQTEILSALHHTCA - SLGP - LRQQC	102
<i>Sorghum_bicolor</i> SORBL_3008G032600/1-247	74 NFMSEAVNYLSEKQTDQKVMFELHDAACS - KSFS - FEEKC	110
<i>Sorghum_bicolor</i> SORBL_3003G055700/1-227	57 QYSTEALFYLTQNETQTEILSILHHECA - SLAP - LKQCC	93
<i>Zea_mays</i> ZEAAMMB73_Zm00001d042734/1-240	63 NFTSEAVTYLGEKQTDRIVEFLHDAACS - QSFS - FEQKC	99
<i>Zea_mays</i> ZEAAMMB73_Zm00001d039719/1-229	59 QYSEALLYLTQNETQTEILSILHHECA - SLAP - LKQCC	95
<i>Aquilegia_coerulea</i> AQUCO_00400489v1/1-223	56 QYSSLAINYLSENKTQTEILDTLHQTCS - RMHG - FKEQC	88
<i>Spinacia_oleracea</i> SOVF_050110/1-231	52 EYTTLAVDYLSQNKTDDEIMESLHKACM - QMHG - LAQCC	92
<i>Helianthus_annuus</i> HannXRQ_Chr10g028629/1-181	13 EYTSEAL IYLQNNKTQEEI I SILHDSCS - KLHS - LSKQC	49
<i>Cynara_cardunculus_var_scolymus</i> Ccd_003008/1-231	63 EYASEALFYLQNNKTQEEI I FILHESCS - KLRS - LEGQC	99
<i>Lactuca_sativa</i> LSAT_9X3806/1-229	64 EYASEALFYLENNKTQKEI I SALHESCD - KLQS - LKKQC	100
<i>Coffea_canephora</i> GSCOC_70002323400/1-294	119 EFAVEAVNYFANNKTQTEILEILYKTC - KMHT - FKQCC	155
<i>Nicotiana_tabacum</i> LOC107812754/1-245	68 EYTAKALKYMANYSKTQTEI IDHLHESCL - KMSF - YKQEC	104
<i>Nicotiana_tabacum</i> LOC107792809/1-238	64 EYTASALGYLSNNETQTKI LDLLLNTCS - KMP I - YKLKC	100
<i>Solanum_tuberosum</i> 102602502/1-242	68 EYTAKALNYMDNNKTQTEI IDRLHKSCS - KMR F - YKEEC	104
<i>Solanum_lycopersicum</i> NA A0A3Q7010/1-238	65 EYTAKALNYMANNKTQTEI IDRLHKSCS - KMR F - YKEEC	101
<i>Cicer_arietinum</i> LOC101491522/1-279	91 EYTSKALDYLNENKTQTEI IDILHNHTCH - QLHT - FEKKC	127
<i>Cicer_arietinum</i> LOC101508260/1-215	41 EYTTKALNYI KDNNTQAEI IDGLHNHTCY - QLLS - FKKQC	77
<i>Medicago_truncatula</i> MTR_7g072560/1-242	52 EYTTKALDYINENKTQSEI IDILHNHTCH - QLHT - FEKKC	88
<i>Medicago_truncatula</i> MtrunA17_Chr4g001314/1-223	49 EYTT EILDY LKDNKTQAKI IDDLHNHTCH - QLPA - FSEQC	85
<i>Medicago_truncatula</i> MTR_029040/1-215	41 EYTT EILDY LKDNKTQAKI IDDLHNHTCH - QLPA - FSEQC	77
<i>Trifoliumpratense</i> L195_g026334/1-194	20 EYTT EIVLDY LKDNNTQAEI IDSLHNHTCH - HLLS - FNQCC	56
<i>Lotus_japonicus</i> NA 3S9R9/1-216	41 EYTT EIVLDY LKDNKTQSEI IDALHNHTCN - QLFS - FKQCC	77
<i>Phaseolus_vulgaris</i> PHAVU_008G084800g/1-222	54 EYTTKALEYI KQNMTEEEI IDTLHNHTCH - LLPS - FKQCC	90
<i>Phaseolus_vulgaris</i> PHAVU_008G0847000g/1-217	45 EYTTAKALNYLGNKTQTEI IDILHNHTCH - QLPR - LHKQC	81
<i>Glycine_max</i> GLYMA_09G277100/1-237	45 EYSTKVL DYLNENKTQTEI IDILHNHTCH - QTSS - FKQCC	81
<i>Glycine_max</i> GLYMA_01G131400/1-216	42 EYTTAEALDY LNDKENQREI IDSLHNHTCN - HILS - FKQCC	78
<i>Eucalyptus_grandis</i> EUGRSUZ_K01273/1-227	52 EYAEAL KYLAENKTQSEI VELLHLTCS - QVPV - FKAEC	88
<i>Eucalyptus_grandis</i> EUGRSUZ_A00687/1-219	45 EYTAQALDY IGDNKTQTEI LELLHKSCS - HLAS - FEQEC	81
<i>Gossypium_hirsutum</i> LOC107896756/1-233	69 EFATEA INFLSQNKQTQTEI IVEVLHKSCS - RIPS - FEQCC	105
<i>Gossypium_hirsutum</i> LOC107935966/1-227	63 EFTTEAVDYLSQNKQTQTEI IELHKSCS - RLRA - FEPQC	99
<i>Gossypium_tomentosum</i> ES332_D10G139500v1/1-233	69 EFATEA IDFLSQNKQTQTEI IVEVLHKSCS - RIPS - FEQCC	105
<i>Gossypium_tomentosum</i> ES332_A02G005200v1/1-227	63 EFTTEAVDYLSQNKQTQTEI IELHKSCS - RLRA - FEPQC	99
<i>Theobroma_cacao</i> TCM_019744/1-228	63 EFANEA IDYLSQNKQTQTEI IVEMLHKSCS - RVPS - FKQCC	99
<i>Brassica_rapa_subsp_pekinensis</i> NA M4D8N0/1-215	45 EYVTSALTYLEKNKTQTEI LEDLHDCS - LIRG - FEQCC	81
<i>Brassica_rapa_subsp_pekinensis</i> NA M4CRM9/1-214	46 EYVNVA I SYLENNQTQQA I EDLHDCS - HMRG - FAQCC	82
<i>Brassica_oleracea_var_oleracea</i> NA A0A0D3DSC3/1-229	61 EYVNVA I SYLENNQTQQA I EDLHDCS - HMRG - FAQCC	97
<i>Brassica_oleracea_var_oleracea</i> NA A0A0D3D313/1-216	46 EYVTSALTYLEKNKTQTEI LEDLHDCS - LIRG - FEQCC	82
<i>Arabidopsis_lyrata_subsp_lyrata</i> ARALYDRAFT_486888/1-220	47 KYVTLA IDY LQDYDYNQNALVEALHISCS - QIPP - LKKQC	83
<i>Arabidopsis_lyrata_subsp_lyrata</i> ARALYDRAFT_66600/1-213	45 EYVTDALSYLEKNVTQAEI IEDLHDCS - QLRG - FSQCC	81
<i>Arabidopsis_thaliana</i> At5g01800/1-217	44 KYVTLV IDY LQDYDYNQNELVEALHISCS - QIPP - LKKQC	80
<i>Arabidopsis_thaliana</i> At3g51730/1-213	45 EYVTDALSYLEKNVTQAEI IEDLHDCS - QLRG - YSQCC	81
<i>Rosa_chinensis</i> RchiOBHm_Chr3g046096/1-229	54 EFASQALDY I SENKTQTEI IAILHNHTCS - QLKS - FSQCC	90
<i>Prunus_persica</i> PRUPE_6G290000/1-253	65 EFAAQALDYLSENKTQTEI IEALHQTCH - QLGS - FKQCC	101
<i>Malus_domestica</i> DVH24_036312/1-296	126 EFADQALDYLNENKTQTEI IEYLHQTCH - QLRS - FNQCC	162
<i>Populus_trichocarpa</i> POPTR_016G133400/1-242	71 EFASQALDYLAENKTQTEI LEKLHRS - RLTT - FEQCC	107
<i>Populus_trichocarpa</i> POPTR_006G107300/1-242	71 EFAAQALDYMAENKTQTEI LEILHKTCS - RLTT - FKQCC	107
<i>Cucumis_sativus</i> Csa_4G331080/1-233	58 SLISQAVEYFADNQQTSEI IGLLRQTG - VAGV - FKEEC	94
<i>Cucumis_melo_var_makuwa</i> E5676_scaffold127G001120/1-249	74 SLISQAVEYFADNQQTSEI IGLLRQTG - VAGV - FKEEC	110
<i>Cucumis_melo_var_makuwa</i> E6C27_scaffold1166G00310/1-233	58 SLISQAVEYFADNQQTSEI IGLLRQTG - VAGV - FKEEC	94



Amborella trichopoda[AMTR_s00007p00225690/1-214 92 TSLVYYLL - - P L F S E I A M I N P E G L C A K V N L C N S E A - N V 127
Amborella trichopoda[AMTR_s00062p00198130/1-320 138 EESF K S Y I E K A V V F L Q - E Y L S G E R L C N S T G L C P G Y G E T I 175
Cinnamomum micranthum f. kanehirae[CKAN_01065200/1-278 94 LETS D T Y M Q Q A I S F L E - D Y F S E K K F C N S T G L C H E N I E T V 131
Cinnamomum micranthum f. kanehirae[CKAN_00757300/1-212 97 DALVYYAA - - P L F F V E I A M I K P E D F C K K V N L C E D A G - F I 132
Anthurium amnicola[Psapl1_1/1-288 111 I K M V E R Y K Y E A V M L L Q - E V V R E D K F C N S T G F C P N N P Q E S 148
Anthurium amnicola[Psapl1_2/1-273 100 TQMA K Y V R Q A S L C V Q - A F L F G E N V C R N I K L C N S S I S R I 137
Zostera marina[ZOSMA_381G00120/1-242 105 V D T A S Y I E E I I S Y L Q - D L F E E E N L C Y D T G L C T E N A - E I 141
Zostera marina[ZOSMA_56G01350/1-232 102 I L L V O Y Y A - - P F F F L E I E T L D P K K F C T K V N L C G A S S - Y I 137
Dendrobium catenatum[MA16_Dca011512/1-222 85 L I L V O Y Y S - - T L F F T S I G K I R P E E F C G R V G L C E A S S - V I 119
Dendrobium catenatum[MA16_Dca020165/1-215 82 N L L M O Y Y A - - P L F F V E I A K W Q P K L L C E K V N L C K E M S - V I 117
Oryza sativa subsp. indica[Osl_34843/1-245 109 L E T M O S Y A - - T L V F A K I A E I K P D E F C K Q Y G L C R D M A - L L 144
Oryza sativa subsp. indica[Osl_19500/1-223 89 T K L V O Y Y V - - S L F F T K V T S L S P E E F C E S V S L C H K V T - F I 124
Oryza sativa subsp. japonica[Os12g0112200/1-245 109 L E T M O S Y A - - T L V F A K I A E I K P D E F C K Q Y G L C R D M A - L L 144
Oryza sativa subsp. japonica[Os05g0334400/1-223 89 T K L V O Y Y V - - S L F F T K V T S L S P E E F C E S V S L C H K V T - F I 124
Brachypodium distachyon[BRADL_4g25580v3/1-245 107 V E L V O S Y A - - T L L F A K I A E I K P D E F C K Q H G L C R D T A - L L 142
Brachypodium distachyon[BRADL_2g04110v3/1-235 102 I T L V O Y Y I - - P L F F M E V S A N P E V F C E S V H L C P K G T - R S 137
Hordeum vulgare subsp. vulgare[NA|F2DBE9/1-246 110 L E M V O S Y A - - T L L F A K I T E I K P D E F C K Q Y G L C R D V S - F L 145
Hordeum vulgare subsp. vulgare[NA|A0A287KA24/1-238 105 M T L V O Y Y I - - P A F F L E V S V L K P E E L C E S A H L C P K G A - A A 140
Triticum aestivum[NA|A0A3B6MM75/1-246 110 L E M V O S Y A - - T L L F A K V T E I K P D E F C K R Y G L C R D V S - F L 145
Triticum aestivum[NA|A0A3B6FV9/1-233 103 I T L V O Y Y I - - P A F F L E V S V V K P E E L C E S A H L C P K G A - A T 138
Sorghum bicolor[SORBI_3008G032600/1-247 111 V E L M O S Y A - - T L L F A K I T E I E P A F C K Q Y G L C R N T A - L F 146
Sorghum bicolor[SORBI_3003G055700/1-227 94 I T L V O Y Y I - - P L F F F E V S M V T P E K F C E S M H L C K N G M - K I 129
Zea mays[ZEA MMB73_Zm00001d042734/1-240 100 V E L M O S Y A - - T L L F A K I T E I R P E A F C K R Y G L C R D T A - L L 135
Zea mays[ZEA MMB73_Zm00001d039719/1-229 96 I T L V O Y Y V - - P L F F L E V S M V T P E K F C E S M H L C K K G M - K I 131
Aquilegia coerulea[AQUCO_00400489v1/1-223 89 T T L V O Y Y A - - P L F F L E V A G V E P V E F C S K M N L C D - - - - I 120
Spinacia oleracea[SOVF_050110/1-231 93 T T L V O Y Y A - - P L F F L E V S S V E P E G F C K K V D L C R N T M - V S 128
Helianthus annuus[HannXRQ_Ch10g028629/1-181 50 I T L V O Y Y A - - P L F F L E L S N I Q P E D F C G K V N L C K E V V - A Y 85
Cynara cardunculus var. scolymus[Cord_003008/1-231 100 T T L V O Y Y A - - P L F F L E L S T I K P S D F C G K V N L C N E V V - A Y 135
Lactuca sativa[LSAT_IX38061/1-229 101 I T L V O Y Y A - - P L F F L E I S T V K P E D F C G K V G L C K E I V - A Y 136
Coffea canephora[GSCOC_T0002323400/1-294 156 T S L V O Y Y A - - P L F F L E I S S V Q P K D F C Q K V D L C E D I V - S I 191
Nicotiana tabacum[LOC107812754/1-245 105 V V L V O Y Y A - - P L F F S E I N K I R P E D F C E K F D L C E R V V - T V 140
Nicotiana tabacum[LOC107792809/1-238 101 T A L V O Q Y V - - R F F L V I S T I K P D D I C Q K V D L C Q K V V - S I 136
Solanum tuberosum[102602502/1-242 105 A I L V O Y Y A - - P L F F L Q I N K M K P E N F C Q Q F G L C E Q V V - I I 140
Solanum lycopersicum[NA|A0A3Q7010/1-238 102 A I L V O Y Y A - - P L F F L Q I N K M K P E N F C Q Q F G L C E Q V V - I I 137
Cicer arietinum[LOC101491522/1-279 128 I N L V O Y Y V - - P L F F L E A T S V Q P G D F C N K V N L C Q T I A - D L 163
Cicer arietinum[LOC101508260/1-215 78 I E L V O Y Y A - - P L F F S K I A I K P G E L C E K F N L C E S A K - A A 112
Medicago truncatula[MTR_7g072560/1-242 89 I N L V O Y Y L - - P L F F S E M S V Q P G D F C N K V N L C Q N I A - N I 124
Medicago truncatula[MtrunA17_Ch4g001314/1-223 86 F E L V O H H V - - Q L F F S K I A R M M P A E L C E K Y H L C E S A T - - I 120
Medicago truncatula[MTR_029040/1-215 78 F E L V O H H V - - Q L F F S K I A R M M P A E L C E K Y H L C E S A T - - I 112
Trifolium pratense[L195_g026334/1-194 57 L K L V O Y Y V - - P L F F S E I A R I N P G E L C D K F N L C E S A K - N Y 92
Lotus japonicus[NA|J359R9/1-216 78 I E L V O N Y V - - P L F F I E L A S V Q P E E L C K T V F L C Q S A K - L I 112
Phaseolus vulgaris[PHAVU_008G084800g/1-222 91 I A L V O Y Y T - - P I F L S E V A S L K P R E F C H K I D I C Q L T E - H I 126
Phaseolus vulgaris[PHAVU_008G0847000g/1-217 82 I I L V O Y Y A - - P L F F L E M A T I Q P E D F C N K I N I C H L I S - Y I 117
Glycine max[GLYMA_09G277100/1-237 82 I T L V O Y Y A - - P L F F L E I V T I Q P G E F C H K V N L C Q L I T - Y I 117
Glycine max[GLYMA_01G131400/1-216 79 I E L V O H Y A - - P L F F L E I A S V Q P G E L C K Q I H I C Q S A K - I I 113
Eucalyptus grandis[EUGRSUZ_K01273/1-227 89 V V L V O Y Y A - - P I F F L E L S T I Q P E E L C K D I S A C K L A A - R V 124
Eucalyptus grandis[EUGRSUZ_A00687/1-219 82 L S L V O S Y A - - T L F F S E V S S V E P E E F C R K V N L C E K K V - F L 117
Gossypium hirsutum[LOC107896756/1-233 106 I T L V O Y Y V - - P L F F V E I S L I Q P E V L C K E V N L C Q K F A - L I 141
Gossypium hirsutum[LOC107935966/1-227 100 I T L V O Y Y A - - P L F F L E I Y S V Q P Q D F C T K F N L C Q K V A - L I 135
Gossypium tomentosum[ES332_D10G139500v1/1-233 106 I T L L O Y Y V - - P L F F V E I S S V Q P E V L C K E V N L C Q K F A - L I 141
Gossypium tomentosum[ES332_A02G005200v1/1-227 100 I T L V O Y Y A - - P L F F L E I S S V Q P Q D F C T K F N L C Q K V A - L I 135
Theobroma cacao[TCM_019744/1-228 100 I T L V O Y Y V - - P L F F M E V S S I R P E D F C Q K V N L C Q K V A - L I 135
Brassica rapa subsp. pekinensis[NA|M4D8N0/1-215 82 I T L V O Y Y L - - P L F F L H L E S F Q P H Y F C K R M N L C G H V V - A L 117
Brassica rapa subsp. pekinensis[NA|M4CRM9/1-214 83 V T L V O Y Y V - - P L F F I Q L E S F Q P Q D F C K R M N L C D K V A - A L 118
Brassica oleracea var. oleracea[NA|A0A0D3DSC3/1-229 98 V T L V O Y Y V - - P L F F I Q L E S F Q P Q D F C K R M N L C D K V A - A L 133
Brassica oleracea var. oleracea[NA|A0A0D3D313/1-216 83 I T L V O Y Y L - - P L F F L H L E S F Q P H Y F C K R M N L C G H V V - A L 118
Arabidopsis lyrata subsp. lyrata[ARALYDRAFT_486888/1-220 84 L S M V O H Y T - - Q L F F T Q V S T I T S D Q I C K R L N L C Q A A T P P F 120
Arabidopsis lyrata subsp. lyrata[ARALYDRAFT_666001/1-213 82 I T L V O Y Y V - - P L F F L Q L E S F Q P H Y F C K R M N L C G K V V - A L 117
Arabidopsis thaliana[At5g01800/1-217 81 L S M V O H Y T - - Q L F F T Q V S T I K S D Q I C K R L N L C Q A V T P A F 117
Arabidopsis thaliana[At3g51730/1-213 82 I S L V O Y Y V - - P L F F L Q L E S F Q P H Y F C K R M N L C G K V V - A L 117
Rosa chinensis[RchiOBHm_Ch3g046096/1-229 91 V T L V O Y Y A - - P L F F L E V T S V E P V D F C R K V N L C Q Q V A - T F 126
Prunus persica[PRUPE_6G29000/1-253 102 I T L V O Y Y A - - P L F F L E V S S L Q P S E F C R K V N L C Q Q V A - L F 137
Malus domestica[DVH24_036312/1-296 163 V T L V O Y Y A - - P L F F L E A T S L Q P S E F C R K V N L C Q Q V A - L F 198
Populus trichocarpa[POPTR_016G133400/1-242 108 I T L V O Y Y S - - S I F F S Y A S S V Q S E D F C R K F N L C Q E M K - T F 143
Populus trichocarpa[POPTR_006G107300/1-242 108 I T L V O Y Y S - - S I F F S Y V S S V Q S D D F C R K Y N L C H E M E - I F 143
Cucumis sativus[Csa_4G331080/1-233 95 I S L V O S Y V - - P L F F S K I S S I E P S S I C Q S A H I C E Q V T - I I 130
Cucumis melo var. makuwa[E5676_scaffold127G001120/1-249 111 I S L V O S Y V - - P L F F S E I S S I E P S S I C Q S A H F C Q V T - I I 146
Cucumis melo var. makuwa[E6C27_scaffold1166G00310/1-233 95 I S L V O S Y V - - P L F F S E I S S I E P S S I C Q S A H F C E Q V T - I I 130

Conservation



Quality



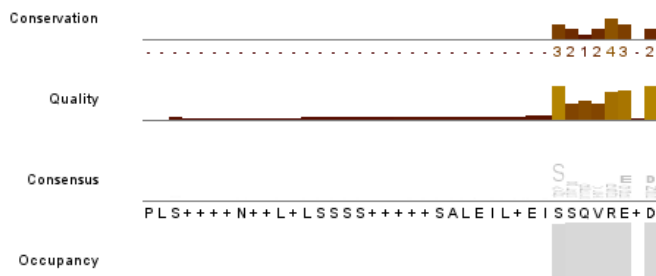
Consensus

LVDYY PLFFLEISSIQPEEFCKKVNLCCQKVA++I
 I T L V D Y Y A E + P L F F L E I S S I Q P E E F C K K V N L C Q K V A + + I

Occupancy



<i>Amborella trichopoda</i> AMTR_s00007p00225690/1-214	128	APKASYLD	135
<i>Amborella trichopoda</i> AMTR_s00062p00198130/1-320	176	--SNNERNIQLNFGSFYNKLSMLRETLLLEVANMAYQ-R	211	
<i>Cinnamomum micranthum_f_kanehirae</i> CKAN_01065200/1-278	132	--S-----LSSIWTTKSLSALEIVNEIMNEVWK-S	158	
<i>Cinnamomum micranthum_f_kanehirae</i> CKAN_00757300/1-212	133	SPQIYG-D	139
<i>Anthurium amnicola</i> Psapl_1/1-288	149	--F-----LSTSSKANLLSAMEHFNEMLSKMQK-D	175	
<i>Anthurium amnicola</i> Psapl_2/1-273	138	PLSTSGENEMLF TMESSEYILSKVLKGLLTISLKHKEYK	176	
<i>Zostera marina</i> ZOSMA_381G00120/1-242	142	--S-----QITHASKIQPLLMPLPESDTN	162	
<i>Zostera marina</i> ZOSMA_56G01350/1-232	138	PRMAVD-D	144
<i>Dendrobium catenatum</i> MA16_Dca011512/1-222	120	SMAKND-E	126
<i>Dendrobium catenatum</i> MA16_Dca020165/1-215	118	HLRKP-D	124
<i>Oryza sativa_subsp_indica</i> Osl_34843/1-245	145	SAVKSE-S	151
<i>Oryza sativa_subsp_indica</i> Osl_19500/1-223	125	RLPRHE-D	131
<i>Oryza sativa_subsp_japonica</i> Os12g0112200/1-245	145	SAMKSE-S	151
<i>Oryza sativa_subsp_japonica</i> Os05g0334400/1-223	125	RLPRHE-D	131
<i>Brachypodium distachyon</i> BRADL_4g25580v3/1-245	143	SISGVKSE-S	151
<i>Brachypodium distachyon</i> BRADL_2g04110v3/1-235	138	RLPTRR-D	144
<i>Hordeum vulgare_subsp_vulgare</i> NA F2DBE9/1-246	146	SLAKSE-S	152
<i>Hordeum vulgare_subsp_vulgare</i> NA A0A287KA24/1-238	141	RSSTRG-E	147
<i>Triticum aestivum</i> NA A0A3B6MMT5/1-246	146	SVAKSE-S	152
<i>Triticum aestivum</i> NA A0A3B6FHG9/1-233	139	RSSTRG-E	146
<i>Sorghum bicolor</i> SORBL_3008G032600/1-247	147	SGVRSN-S	153
<i>Sorghum bicolor</i> SORBL_3003G055700/1-227	130	SLPTRE-G	136
<i>Zea mays</i> ZEAMMB73_Zm00001d042734/1-240	136	SGVGSD-S	142
<i>Zea mays</i> ZEAMMB73_Zm00001d039719/1-229	132	SLPTRE-G	138
<i>Aquilegia coerulea</i> AQUOCO_00400489v1/1-223	121	MESSPQ-D	127
<i>Spinacia oleracea</i> SOVF_050110/1-231	129	SMPEKR-N	135
<i>Helianthus annuus</i> HannXRG_Chr10g028629/1-181	86	ARELSE-N	92
<i>Cynara cardunculus_var_scolymus</i> Ccrd_003008/1-231	136	AQESQ-N	142
<i>Lactuca sativa</i> LSAT_9X38061/1-229	137	AHEFSE-N	143
<i>Coffea canephora</i> GSCOC_T0002323400/1-294	192	SQSLSK-N	198
<i>Nicotiana tabacum</i> LOC107812754/1-245	141	SQVLSG-Q	147
<i>Nicotiana tabacum</i> LOC107792809/1-238	137	SQQFSQ-N	143
<i>Solanum tuberosum</i> 102602502/1-242	141	SQVLSG-K	147
<i>Solanum lycopersicum</i> NA A0A3Q700/1-238	138	SQALSG-K	144
<i>Cicer arietinum</i> LOC101491522/1-279	164	SLQVQE-N	170
<i>Cicer arietinum</i> LOC101508260/1-215	113	SSQVQG-N	119
<i>Medicago truncatula</i> MTR_7g072560/1-242	125	SLKVQE-N	131
<i>Medicago truncatula</i> MtrunA17_Chr4g001314/1-223	121	SSQVHG-N	127
<i>Medicago truncatula</i> MTR_029040/1-215	113	SSQVHG-N	119
<i>Trifolium pratense</i> L195_g026334/1-194	93	ARVRG-N	98
<i>Lotus japonicus</i> NA 3S9R9/1-216	113	SSQVRE-N	119
<i>Phaseolus vulgaris</i> PHAVU_008G084800g/1-222	127	SLQVQE-D	133
<i>Phaseolus vulgaris</i> PHAVU_008G0847000g/1-217	118	SSQVQE-D	124
<i>Glycine max</i> GLYMA_09G277100/1-237	118	SLLVQE-D	124
<i>Glycine max</i> GLYMA_01G131400/1-216	114	SSEVEG-N	120
<i>Eucalyptus grandis</i> EUGRSUZ_K01273/1-227	125	SPKLKE-D	131
<i>Eucalyptus grandis</i> EUGRSUZ_A00687/1-219	118	SSQLQE-D	124
<i>Gossypium hirsutum</i> LOC107896756/1-233	142	STQIRE-D	148
<i>Gossypium hirsutum</i> LOC107935966/1-227	136	SSQFRE-D	142
<i>Gossypium tomentosum</i> ES332_D10G139500v1/1-233	142	STQIRE-D	148
<i>Gossypium tomentosum</i> ES332_A02G005200v1/1-227	136	SSQFRE-D	142
<i>Theobroma cacao</i> TCM_019744/1-228	136	STQIRE-D	142
<i>Brassica rapa_subsp_pekinensis</i> NA M4DBN0/1-215	118	VKEARQ-D	124
<i>Brassica rapa_subsp_pekinensis</i> NA M4CRM9/1-214	119	VEEARQ-D	125
<i>Brassica oleracea_var_oleracea</i> NA A0A0D3DSC3/1-229	134	VEEARQ-D	140
<i>Brassica oleracea_var_oleracea</i> NA A0A0D3D313/1-216	119	VQEARQ-D	125
<i>Arabidopsis lyrata_subsp_lyrata</i> ARALYDRAFT_486888/1-220	121	ASQVHQ-G	127
<i>Arabidopsis lyrata_subsp_lyrata</i> ARALYDRAFT_666001/1-213	118	VEEVQR-D	124
<i>Arabidopsis thaliana</i> At5g01800/1-217	118	ASQVHQ-G	124
<i>Arabidopsis thaliana</i> At3g51730/1-213	118	VEEARQ-D	124
<i>Rosa chinensis</i> RchiOBHm_Chr3g0460961/1-229	127	SSQLRE-D	133
<i>Prunus persica</i> PRUPE_6G290000/1-253	138	SSQLRE-D	144
<i>Malus domestica</i> DVH24_036312/1-296	199	SSQFKE-D	205
<i>Populus trichocarpa</i> POPTR_016G133400/1-242	144	SAKRND-D	150
<i>Populus trichocarpa</i> POPTR_006G107300/1-242	144	SAKHQE-D	150
<i>Cucumis sativus</i> Csa_4G331080/1-233	131	SSLFQD-H	137
<i>Cucumis melo_var_makuwa</i> E5676_scaffold127G001120/1-249	147	SSLVQD-H	153
<i>Cucumis melo_var_makuwa</i> E6C27_scaffold1166G00310/1-233	131	SSLFQD-H	137



<i>Amborella trichopoda</i> AMTR_s00007p00225690/1-214	136	- - - - -	G	C	E	F	C	140
<i>Amborella trichopoda</i> AMTR_s00062p00198130/1-320	212	- - - - -	Q	E	Q	S	E	244
<i>Cinnamomum micranthum</i> _f_kanehirae CKAN_01065200/1-278	159	- - - - -	L	A	E	K	E	197
<i>Cinnamomum micranthum</i> _f_kanehirae CKAN_00757300/1-212	140	- - - - -	S	S	S	V	C	144
<i>Anthurium amnicola</i> Psapl_1_1/1-288	176	- - - - -	L	P	M	E	-	189
<i>Anthurium amnicola</i> Psapl_1_2/1-273	177	- - - - -	S	C	A	P	-	181
<i>Zostera marina</i> ZOSMA_381G00120/1-242	163	- - - - -	V	C	T	V	-	167
<i>Zostera marina</i> ZOSMA_56G01350/1-232	145	- - - - -	T	G	I	C	-	149
<i>Dendrobium catenatum</i> MA16_Dca011512/1-222	127	- - - - -	K	C	A	L	-	131
<i>Dendrobium catenatum</i> MA16_Dca020165/1-215	125	- - - - -	P	C	T	L	-	129
<i>Oryza sativa</i> _subsp_indica Osl_34843/1-245	152	- - - - -	T	C	V	F	-	156
<i>Oryza sativa</i> _subsp_indica Osl_19500/1-223	132	- - - - -	S	C	D	L	-	136
<i>Oryza sativa</i> _subsp_japonica Os12g0112200/1-245	152	- - - - -	T	C	V	F	-	156
<i>Oryza sativa</i> _subsp_japonica Os05g0334400/1-223	132	- - - - -	S	C	D	L	-	136
<i>Brachypodium distachyon</i> BRADL_4g25580v3/1-245	152	- - - - -	T	C	V	F	-	156
<i>Brachypodium distachyon</i> BRADL_2g04110v3/1-235	145	- - - - -	T	C	G	L	-	149
<i>Hordeum vulgare</i> _subsp_vulgare NA F2DBE9/1-246	153	- - - - -	T	C	A	F	-	157
<i>Hordeum vulgare</i> _subsp_vulgare NA A0A287KA24/1-238	148	- - - - -	A	C	G	L	-	152
<i>Triticum aestivum</i> NA A0A3B6MMT5/1-246	153	- - - - -	T	C	A	F	-	157
<i>Triticum aestivum</i> NA A0A3B6FHG9/1-233	146	- - - - -	A	C	G	L	-	150
<i>Sorghum bicolor</i> SORBI_3008G032600/1-247	154	- - - - -	T	C	V	F	-	158
<i>Sorghum bicolor</i> SORBI_3003G055700/1-227	137	- - - - -	T	C	G	L	-	141
<i>Zea mays</i> ZEMMB73_Zm00001d042734/1-240	143	- - - - -	T	C	V	F	-	147
<i>Zea mays</i> ZEMMB73_Zm00001d039719/1-229	139	- - - - -	T	C	G	L	-	143
<i>Aquilegia coerulea</i> AQUCO_00400489v1/1-223	128	- - - - -	S	C	T	V	-	132
<i>Spinacia oleracea</i> SOVF_050110/1-231	136	- - - - -	K	D	L	C	-	140
<i>Helianthus annuus</i> HannXRQ_Chr10g028629/1-181	93	- - - - -	S	C	D	V	-	97
<i>Cynara cardunculus</i> _var_scolymus Cord_003008/1-231	143	- - - - -	S	C	D	V	-	147
<i>Lactuca sativa</i> LSAT_9X38061/1-229	144	- - - - -	S	C	D	V	-	148
<i>Coffea canephora</i> GSCOC_T0002323400/1-294	199	- - - - -	S	C	E	L	-	203
<i>Nicotiana tabacum</i> LOC107812754/1-245	148	- - - - -	S	C	D	L	-	152
<i>Nicotiana tabacum</i> LOC107792809/1-238	144	- - - - -	G	D	L	C	-	148
<i>Solanum tuberosum</i> 102602502/1-242	148	- - - - -	N	D	L	C	-	152
<i>Solanum lycopersicum</i> NA A0A3Q700/1-238	145	- - - - -	N	C	N	L	-	149
<i>Cicer arietinum</i> LOC101491522/1-279	171	- - - - -	S	C	E	F	-	175
<i>Cicer arietinum</i> LOC101508260/1-215	120	- - - - -	S	C	G	L	-	124
<i>Medicago truncatula</i> MTR_7g072560/1-242	132	- - - - -	T	C	E	F	-	136
<i>Medicago truncatula</i> MtrunA17_Chr4g001314/1-223	128	- - - - -	S	C	G	F	-	132
<i>Medicago truncatula</i> MTR_029040/1-215	120	- - - - -	S	C	G	F	-	124
<i>Trifolium pratense</i> L195_g026334/1-194	99	- - - - -	S	C	G	F	-	103
<i>Lotus japonicus</i> NA J3S9R9/1-216	120	- - - - -	S	C	G	F	-	124
<i>Phaseolus vulgaris</i> PHAVU_008G084800g/1-222	134	- - - - -	A	C	E	F	-	138
<i>Phaseolus vulgaris</i> PHAVU_008G0847000g/1-217	125	- - - - -	S	C	G	F	-	129
<i>Glycine max</i> GLYMA_09G277100/1-237	125	- - - - -	T	S	G	F	-	129
<i>Glycine max</i> GLYMA_01G131400/1-216	121	- - - - -	S	C	D	S	-	125
<i>Eucalyptus grandis</i> EUGRSUZ_K01273/1-227	132	- - - - -	S	C	E	F	-	136
<i>Eucalyptus grandis</i> EUGRSUZ_A00687/1-219	125	- - - - -	S	C	E	L	-	129
<i>Gossypium hirsutum</i> LOC107896756/1-233	149	- - - - -	C	G	V	-	-	153
<i>Gossypium hirsutum</i> LOC107935966/1-227	143	- - - - -	S	C	G	M	-	147
<i>Gossypium tomentosum</i> ES332_D10G139500v1/1-233	149	- - - - -	C	G	V	-	-	153
<i>Gossypium tomentosum</i> ES332_A02G005200v1/1-227	143	- - - - -	S	C	G	M	-	147
<i>Theobroma cacao</i> TCM_019744/1-228	143	- - - - -	S	C	G	M	-	147
<i>Brassica rapa</i> _subsp_pekinensis NA M4DBNO/1-215	125	- - - - -	T	C	G	V	-	129
<i>Brassica rapa</i> _subsp_pekinensis NA M4CRM9/1-214	126	- - - - -	S	C	A	V	-	130
<i>Brassica oleracea</i> _var_oleracea NA A0A0D3DSC3/1-229	141	- - - - -	S	C	A	V	-	145
<i>Brassica oleracea</i> _var_oleracea NA A0A0D3D313/1-216	126	- - - - -	T	C	D	V	-	130
<i>Arabidopsis lyrata</i> _subsp_lyrata ARALYDRAFT_486888/1-220	128	- - - - -	N	C	E	A	-	132
<i>Arabidopsis lyrata</i> _subsp_lyrata ARALYDRAFT_666001/1-213	125	- - - - -	S	C	G	V	-	129
<i>Arabidopsis thaliana</i> At5g01800/1-217	125	- - - - -	N	C	E	A	-	129
<i>Arabidopsis thaliana</i> At3g51730/1-213	125	- - - - -	S	C	G	V	-	129
<i>Rosa chinensis</i> RchiOBHm_Chr3g0460961/1-229	134	- - - - -	S	C	G	L	-	138
<i>Prunus persica</i> PRUPE_6G290000/1-253	145	- - - - -	S	C	G	L	-	149
<i>Malus domestica</i> DVH24_036312/1-296	206	- - - - -	S	C	G	L	-	210
<i>Populus trichocarpa</i> POPTR_016G133400/1-242	151	- - - - -	S	C	S	I	-	155
<i>Populus trichocarpa</i> POPTR_006G107300/1-242	151	- - - - -	S	C	S	I	-	155
<i>Cucumis sativus</i> Csa_4G331080/1-233	138	- - - - -	N	C	E	F	-	142
<i>Cucumis melo</i> _var_makuwa E5676_scaffold127G001120/1-249	154	- - - - -	N	C	E	F	-	158
<i>Cucumis melo</i> _var_makuwa E6C27_scaffold1166G00310/1-233	138	- - - - -	N	C	E	F	-	142

Conservation



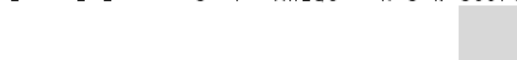
Quality



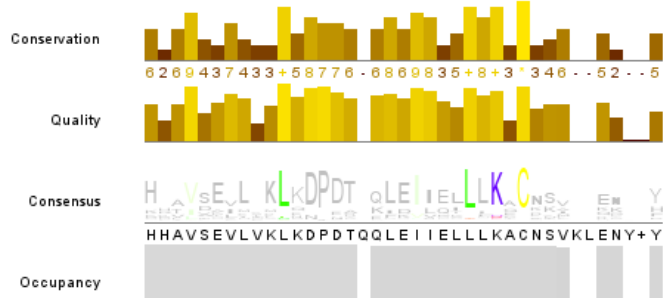
Consensus



Occupancy



<i>Amborella trichopoda</i> [AMTR_s00007p00225690/1-214	141 HRAVLEILMKLKD PPT- QLEIVEILIKKKA- QG- - F 174
<i>Amborella trichopoda</i> [AMTR_s00062p00198130/1-320	245 HRAIDIEIKDRNPVT- KKLKIKILLQACEEV- - QD- - H 278
<i>Cinnamomum micranthum_f_kanehirae</i> [CKAN_01065200/1-278	198 HNALNELLDLENPET- KIKVLKVLKACEEV- - EN- - H 231
<i>Cinnamomum micranthum_f_kanehirae</i> [CKAN_00757300/1-212	145 HEAVVEVLLKLD PPT- - - - - - - - - - - - - - - - - 160
<i>Anthurium amnicola</i> [Psapl_1_1-288	190 HSLVDKIHNALED PDK- QLKFIITSLKACET- - - EA- - F 222
<i>Anthurium amnicola</i> [Psapl_2/1-273	182 RNAMENLHNVGNTER- QIMIKILHACEVA- - DP- - Y 215
<i>Zostera marina</i> [ZOSMA_381G00120/1-242	168 LNATDAFINALDDPEM- KIKVILKLFKACQV- - EN- - H 201
<i>Zostera marina</i> [ZOSMA_56G01350/1-232	150 NDAISEFLVKLED PPT- ELEIEGLNLNKNKV- - EQ- - Y 183
<i>Dendrobium catenatum</i> [MA16_Dca011512/1-222	132 HRVVVGMMLKKNPDA- QLEILQMLFKKCDKM- - KN- - Y 165
<i>Dendrobium catenatum</i> [MA16_Dca020165/1-215	130 HNAINEVLTKLED PPT- QLEVIQILIKACNKA- - EN- - F 163
<i>Oryza sativa_subsp_indica</i> [Osl_34843/1-245	157 HHIIDEIMSKLD PPA- EFEIIQLLLKCNKI- - EG- - H 190
<i>Oryza sativa_subsp_indica</i> [Osl_19500/1-223	137 HEVVDIELTDENPDV- ELKIEVLLKGCNNA- - EN- - F 170
<i>Oryza sativa_subsp_japonica</i> [Osl2g0112200/1-245	157 HHIIDEIMSKLD PPA- EFEIIQLLLKCNKI- - EG- - H 190
<i>Oryza sativa_subsp_japonica</i> [Osl5g0334400/1-223	137 HEVVDIELTDENPDV- ELKIEVLLKGCNNA- - EN- - F 170
<i>Brachypodium distachyon</i> [BRADL_4g25580v3/1-245	157 HHLVDEVMSKLD PPA- EFEIIQLLLKCNKI- - EG- - H 190
<i>Brachypodium distachyon</i> [BRADL_2g04110v3/1-235	150 HHVLVEVLTMLKDPNM- KLEIVGLFLKQCSKA- - EN- - Y 183
<i>Hordeum vulgare_subsp_vulgare</i> [WAF2DBE9/1-246	158 HHLVDVVLKMLKDPDS- QFEILQLLLKCNKV- - HD- - H 191
<i>Hordeum vulgare_subsp_vulgare</i> [WAF2DBE9/1-246	153 HHVLVEVLTMLKDPNT- KLEIVGLFLKQCSKA- - KN- - Y 186
<i>Triticum aestivum</i> [WAF2DBE9/1-246	158 HHLVDVVLKMLKDPDA- QFEILQLLLKCNKV- - KG- - H 191
<i>Triticum aestivum</i> [WAF2DBE9/1-246	151 HHVLVEVLTMLKDPNT- KLEIVGLFLKQCSKA- - KN- - Y 184
<i>Sorghum bicolor</i> [SORBL_3008G032600/1-247	159 HHLDEIMSKLD PPA- EIEIEILIKCNKI- - EG- - H 192
<i>Sorghum bicolor</i> [SORBL_3003G055700/1-227	142 HHVVVEVLVLMKDPNT- QLEVIDLLNKTCSKA- - QN- - Y 175
<i>Zea mays</i> [ZEAAMB73_Zm00001d042734/1-240	148 HHLDEIMSKLD PPA- EFEIIQLLLKCNKI- - EG- - H 181
<i>Zea mays</i> [ZEAAMB73_Zm00001d039719/1-229	144 HHVVVEILIMKDPNM- QLEVIDLLNKTCSKA- - QN- - Y 177
<i>Aquilegia coerulea</i> [AQUQC_00400489v1/1-223	133 HEAVAEILNKLED PPT- QLEIEILLLKVCNKT- - DQ- - Y 166
<i>Spinacia oleracea</i> [SOVF_050110/1-231	141 HRAVDEALEKLD PPT- ELEVIQLLLKACNSV- - GK- - L 174
<i>Helianthus annuus</i> [HannXRQ_Chr10g028629/1-181	98 NLAVERSEILKLD PPDN- QLQILELLLKQCKSV- - EK- - Y 131
<i>Cynara cardunculus_var_scolymus</i> [Cord_003008/1-231	148 HLAVERSEVTLKKNPDA- QLEILELLLKQCKTV- - EK- - Y 181
<i>Lactuca sativa</i> [LSAT_9X3806/1-229	149 HLAVERSEVSLKDPDN- QLEILQLLLKQCKTV- - EK- - Y 182
<i>Coffea canephora</i> [GSCOC_70002323400/1-294	204 HTVVSEAITKLKDPDT- QLEIVEALLKACDAV- - QG- - H 237
<i>Nicotiana tabacum</i> [LOC107812754/1-245	153 HQVVTAEASKLD PPT- QLEILELLLKQCKTV- - EK- - Y 188
<i>Nicotiana tabacum</i> [LOC107792809/1-238	149 HQVVKETLLKKNPDT- QLDILELLLKQCKSV- - EK- - Y 182
<i>Solanum tuberosum</i> [102602502/1-242	153 HKVVTAEASKLD PPT- RLEILELLLKQCKAI- - KP- - Y 186
<i>Solanum lycopersicum</i> [WAF2DBE9/1-246	150 HKVLVDVESKLD PPT- QFEILELLLKQCKAI- - EP- - Y 183
<i>Cicer arietinum</i> [LOC101491522/1-279	176 EDTVSALLAKLD PPT- QLEIEITLLKVCNSV- - DK- - Y 209
<i>Cicer arietinum</i> [LOC101508260/1-215	125 KDAVAALLVELND PPT- KLEIMEKLLKCKSL- - EK- - Y 158
<i>Medicago truncatula</i> [MTR_7g072560/1-242	137 EETVSSLLDKLD PPT- ELEIEILLLKVCNSV- - DK- - F 170
<i>Medicago truncatula</i> [MtrunA17_Chr4g001314/1-223	133 KDTVAELLVELND PPT- KLEIQKLLKACNNM- - EK- - Y 166
<i>Medicago truncatula</i> [MTR_029040/1-215	125 KDTVAELLVELND PPT- KLEIQKLLKACNNM- - EK- - Y 158
<i>Trifolium pratense</i> [L195_g026334/1-194	104 KDTVAQLLVELND PPT- KLEIMEKLLKACNSL- - EN- - H 137
<i>Lotus japonicus</i> [WAF2DBE9/1-246	125 KDAVALLVKLD PPDN- KLEIMESLLKACNSM- - EK- - KL 159
<i>Phaseolus vulgaris</i> [PHAVU_008G084800g/1-222	139 EETVSTLLVKLD PPT- KLEIEITLLKLCNAV- - EK- - Y 172
<i>Phaseolus vulgaris</i> [PHAVU_008G0847000g/1-217	130 KDTVSTLLAKLD PPT- KLEIEITLLKVCNSV- - EK- - Y 163
<i>Glycine_max</i> [GLYMA_09G277100/1-237	130 KDTVSTLLSAKLD PPT- KLEIEITLLKVCNSV- - EK- - Y 163
<i>Glycine_max</i> [GLYMA_01G131400/1-216	126 KDTVSALLVKLD PPT- KLEIMEALLKACNSM- - DK- - F 159
<i>Eucalyptus grandis</i> [EUGRSUZ_K01273/1-227	137 HDAVSQVLDKLD PPT- QMDIIEILLKQCKSL- - ES- - Y 170
<i>Eucalyptus grandis</i> [EUGRSUZ_A00687/1-219	130 HHAVERSEVLDKLD PPT- QMDIIEILLKQCKSL- - EN- - Y 163
<i>Gossypium hirsutum</i> [LOC107896756/1-233	154 HHVPSEVLTMLKDPPT- QLEIEILLKGCDSV- - QN- - Y 187
<i>Gossypium hirsutum</i> [LOC107935966/1-227	148 HRAISEVLMKLD PPT- KLEILELLLKGCNSV- - QN- - Y 181
<i>Gossypium tomentosum</i> [ES332_D10G139500v1/1-233	154 HHAVERSEVLTMLKDPPT- QLEIEILLKGCDSV- - QN- - Y 187
<i>Gossypium tomentosum</i> [ES332_A02G005200v1/1-227	148 HRAISEVLMKLD PPT- KLEILELLLKGCNSV- - QN- - Y 181
<i>Theobroma cacao</i> [TCM_019744/1-228	148 HRAIAEVLILKLD PPTQLDIIEILLKGCDSM- - QN- - Y 182
<i>Brassica rapa_subsp_pekinensis</i> [WAF2DBE9/1-246	131 HRTVSEILIKLD PPT- QLDVIELLLKQCKSF- - KNEYE 165
<i>Brassica rapa_subsp_pekinensis</i> [WAF2DBE9/1-246	130 HKTVSDILIKLD PPT- QLDVIELLLKQCKSF- - KN- - Y 164
<i>Brassica oleracea_var_oleracea</i> [WAF2DBE9/1-246	146 HKTVSDILIKLD PPT- QLDVIELLLKQCKSF- - KN- - H 179
<i>Brassica oleracea_var_oleracea</i> [WAF2DBE9/1-246	131 HRTVSEILIKLD PPT- QLDVIELLLKQCKSF- - KNEYD 166
<i>Arabidopsis lyrata_subsp_lyrata</i> [ARALYDRAFT_486888/1-220	133 RQTVSEVVAKLKDPQT- KLKIIRLLLKQCKSL- - NN- - Y 168
<i>Arabidopsis lyrata_subsp_lyrata</i> [ARALYDRAFT_66600/1-213	130 HRTVSEILIKLD PPT- QLDVIELLLKQCKSL- - KN- - Y 163
<i>Arabidopsis thaliana</i> [At5g01800/1-217	130 RETVSEVVTMLKDPPT- KLKIIRLLLKQCKSL- - NN- - Y 163
<i>Arabidopsis thaliana</i> [At3g51730/1-213	130 HRTVSEILIKLD PPT- QLDVIELLLKQCKSL- - KN- - Y 163
<i>Rosa chinensis</i> [RchiOBHm_Chr3g046096/1-229	139 HRAVERSEVLAKLD PPT- QLEIEILLKACNSV- - EN- - Y 172
<i>Prunus persica</i> [PRUPE_6G290000/1-253	150 HRAVERSEVLKLD PPT- QLEIEILLKACNSV- - EN- - Y 183
<i>Malus domestica</i> [DVH24_036312/1-296	211 HRAVERSEVLKLD PPT- QLEIEILLKACNSV- - EN- - Y 244
<i>Populus trichocarpa</i> [POPTR_016G133400/1-242	156 QRAVERSEVLKLD PPT- QLEIEILLKACNSM- - EK- - Y 189
<i>Populus trichocarpa</i> [POPTR_006G107300/1-242	156 QHAISEVVLKLD PPT- QLEIIDLKACNSM- - EN- - Y 189
<i>Cucumis sativus</i> [Csa_4G331080/1-233	143 HQTISKILDKLD PPT- QIEILQTLNMCDSI- - NY- - R 176
<i>Cucumis_melo_var_makuwa</i> [E5676_scaffold127G001120/1-249	159 HQTISKILDKLD PPT- QIEILQTLNMCDSI- - NY- - R 192
<i>Cucumis_melo_var_makuwa</i> [E6C27_scaffold1166G00310/1-233	143 HQTISKILDKLD PPT- QIEILQTLNMCDSI- - NY- - R 176



<i>Amborella_trichopoda</i> AMTR_s00007p00225690/1-214	175	VEK	C	D	L	V	F	E	Y	A	P	L	L	L	I	N	A	E	Q	F	L	E	T	K	D	I	C	A	S	V	H	V	K	A	F	-	212					
<i>Amborella_trichopoda</i> AMTR_s00062p00198130/1-320	279	VKE	C	K	L	V	L	E	Y	V	P	L	I	L	V	N	L	E	K	Y	L	K	N	N	D	I	C	A	M	L	H	V	C	K	D	H	-	316				
<i>Cinnamomum_micranthum_f_kanehirae</i> CKAN_01065200/1-278	232	VKE	C	K	L	V	F	E	Y	G	P	L	I	L	A	N	V	E	K	F	L	I	A	K	N	D	L	C	S	I	M	H	I	C	N	S	R	-	269			
<i>Cinnamomum_micranthum_f_kanehirae</i> CKAN_00757300/1-212	161	-	-	Q	C	K	L	V	F	E	Y	G	P	L	I	M	A	N	A	E	R	F	L	E	K	H	D	V	C	V	S	L	H	V	C	K	D	S	-	196		
<i>Anthurium_amicola</i> Psapl_1_1-288	223	I	Y	R	C	K	L	V	F	A	Y	G	P	I	V	I	S	N	L	Q	K	-	I	V	S	M	D	L	C	H	M	V	H	L	C	K	D	P	-	259		
<i>Anthurium_amicola</i> Psapl_1_2/1-273	216	VH	Q	C	K	M	V	L	V	Y	G	P	L	L	L	G	N	V	Q	K	I	L	M	D	N	D	L	C	Y	T	M	N	M	C	K	D	P	L	-	254		
<i>Zostera_marina</i> ZOSMA_381G00120/1-242	202	VQ	Q	C	K	L	V	F	E	Y	A	P	I	F	L	S	R	I	E	K	Y	L	K	N	G	E	L	C	T	L	L	Q	L	C	S	M	D	-	239			
<i>Zostera_marina</i> ZOSMA_56G01350/1-232	184	TQ	Q	C	T	L	V	F	E	Y	G	P	L	I	L	T	N	A	G	K	Y	L	Q	N	L	N	I	C	K	L	I	H	A	C	N	E	D	-	221			
<i>Dendrobium_catenatum</i> MA16_Dca011512/1-222	166	AHE	C	K	W	L	V	M	H	Y	G	P	Y	I	L	T	K	G	E	K	F	L	E	T	N	D	V	C	A	S	I	H	A	C	S	S	K	-	203			
<i>Dendrobium_catenatum</i> MA16_Dca020165/1-215	164	AHQ	C	K	L	V	L	E	Y	G	P	I	I	M	A	N	T	Q	K	F	L	E	K	T	D	I	C	T	A	I	H	V	C	K	A	Q	-	201				
<i>Oryza_sativa_subsp_indica</i> Osl_34843/1-245	191	QQ	Q	C	K	R	M	V	L	Q	Y	P	L	V	L	V	N	G	E	K	F	L	E	K	N	D	V	C	A	M	I	Q	A	C	D	A	G	-	228			
<i>Oryza_sativa_subsp_indica</i> Osl_19500/1-223	171	VQ	Q	C	K	L	I	I	Q	N	A	P	I	I	L	E	H	I	K	K	F	L	K	K	R	D	F	C	N	S	I	H	V	C	G	G	K	-	208			
<i>Oryza_sativa_subsp_japonica</i> Osl2g0112200/1-245	191	QQ	Q	C	K	R	M	V	L	Q	Y	P	L	V	L	V	N	G	E	K	F	L	E	K	N	D	V	C	A	M	I	Q	A	C	D	A	G	-	228			
<i>Oryza_sativa_subsp_japonica</i> Osl5g0334400/1-223	171	VQ	Q	C	K	L	I	I	Q	N	A	P	I	I	L	E	H	I	K	K	F	L	K	K	R	D	F	C	N	S	I	H	V	C	G	G	K	-	208			
<i>Brachypodium_distachyon</i> BRADL_4g25580v3/1-245	191	VQ	Q	C	K	R	L	V	L	Q	Y	P	L	I	L	V	N	G	E	K	F	L	E	K	N	D	I	C	T	I	V	Q	A	C	N	T	-	228				
<i>Brachypodium_distachyon</i> BRADL_2g04110v3/1-235	184	AP	Q	C	K	R	L	V	L	E	Y	P	L	I	L	V	K	T	Q	K	L	E	T	T	D	V	C	S	D	I	H	A	C	K	A	V	-	221				
<i>Hordeum_vulgare_subsp_vulgare</i> NA F2DBE9/1-246	192	VQE	C	K	R	M	V	L	E	Y	V	P	L	I	L	V	N	G	E	K	L	E	K	K	D	V	C	T	L	M	Q	A	C	D	A	G	-	229				
<i>Hordeum_vulgare_subsp_vulgare</i> NA A0A287KA24/1-238	187	EP	Q	C	K	R	L	V	L	D	Y	I	P	L	I	L	V	K	T	Q	T	F	E	T	T	D	V	C	F	T	H	A	C	K	T	G	-	224				
<i>Triticum_aestivum</i> NA A0A3B6MMT5/1-246	192	VQE	C	K	R	M	V	L	E	Y	V	P	L	I	L	V	N	G	E	K	L	E	K	K	D	V	C	T	L	M	Q	A	C	D	A	G	-	229				
<i>Triticum_aestivum</i> NA A0A3B6FHG9/1-233	185	EP	Q	C	K	R	L	V	L	D	Y	I	P	L	I	L	V	K	T	Q	N	F	E	T	T	D	V	C	F	A	T	H	A	C	K	T	G	-	222			
<i>Sorghum_bicolor</i> SORBL_3008G032600/1-247	193	VQ	Q	C	K	R	L	V	L	Q	Y	P	L	I	L	V	N	G	E	K	F	L	E	K	N	D	V	C	A	L	V	Q	A	C	P	A	S	-	230			
<i>Sorghum_bicolor</i> SORBL_3003G055700/1-227	176	EQ	Q	C	K	L	V	L	K	Y	I	P	L	I	L	V	K	G	E	K	F	L	E	T	T	D	V	C	S	A	I	H	A	C	K	A	G	-	213			
<i>Zea_mays</i> ZEMMB73_Zm00001d042734/1-240	182	VQ	Q	C	K	R	L	V	L	Q	Y	I	P	L	I	L	V	N	G	E	K	F	L	E	K	N	D	V	C	A	L	A	Q	A	C	P	A	-	219			
<i>Zea_mays</i> ZEMMB73_Zm00001d039719/1-229	178	EQ	Q	C	K	R	L	V	L	K	Y	I	P	L	I	L	V	K	G	Q	K	F	L	E	T	T	D	V	C	S	V	I	H	A	C	K	A	G	-	215		
<i>Aquilegia_coerulea</i> AQUUCO_00400489v1/1-223	167	AK	K	C	K	L	V	F	E	Y	G	P	L	I	M	A	N	A	D	K	F	L	V	K	N	D	L	C	T	A	I	H	A	C	K	T	S	-	204			
<i>Spinacia_oleracea</i> SOVF_050110/1-231	175	T	K	K	C	S	M	V	F	E	F	G	P	L	I	L	V	D	A	G	K	F	I	Q	N	D	L	C	S	T	F	H	A	C	S	R	Y	-	212			
<i>Helianthus_annuus</i> HannXRQ_Chr10g028629/1-181	132	V	P	K	C	K	L	V	F	E	Y	A	P	L	I	L	A	N	A	E	Q	F	L	E	K	E	D	I	C	A	K	L	H	A	C	D	I	N	-	169		
<i>Cynara_cardunculus_var_scolymus</i> Cord_003008/1-231	182	I	P	K	C	T	L	V	F	E	Y	A	P	L	I	L	A	N	A	E	Q	F	L	E	K	E	D	I	C	G	K	L	H	A	C	D	S	Y	-	219		
<i>Lactuca_sativa</i> LSAT_9X3806/1-229	183	L	P	K	C	S	L	V	F	E	Y	A	P	L	I	L	A	N	A	E	Q	F	L	E	K	E	D	I	C	S	K	L	H	A	C	D	S	Y	-	220		
<i>Coffea_canephora</i> GSCOC_70002323400/1-294	238	V	N	K	C	K	R	M	V	F	E	Y	V	P	I	L	V	N	A	E	Q	F	L	E	T	K	D	I	C	T	M	L	H	A	C	E	S	A	-	275		
<i>Nicotiana_tabacum</i> LOC107812754/1-245	189	A	R	K	C	K	L	I	F	E	Y	A	P	V	I	L	V	N	A	E	Q	F	L	E	K	N	D	V	C	A	I	L	H	D	C	E	P	A	-	226		
<i>Nicotiana_tabacum</i> LOC107792809/1-238	183	A	N	K	C	K	M	V	F	E	Y	A	P	V	I	L	V	N	A	E	H	F	L	E	K	N	D	V	C	T	I	L	H	A	C	E	P	A	-	220		
<i>Solanum_tuberosum</i> 102602502/1-242	187	A	K	K	C	K	L	V	F	E	F	A	P	V	I	L	I	N	A	E	Q	F	L	E	Q	N	D	V	C	A	I	L	H	A	C	E	P	A	-	224		
<i>Solanum_lycopersicon</i> NA A0A3Q7I00/1-238	184	T	K	K	C	K	L	V	F	E	F	A	P	V	I	L	V	N	A	E	Q	F	L	E	Q	N	D	V	C	A	I	L	H	A	C	E	P	A	-	221		
<i>Cicer_arietinum</i> LOC10149152/1-279	210	A	S	K	C	K	R	V	V	L	E	Y	G	P	L	V	F	E	N	A	E	K	F	L	E	K	T	D	I	C	T	A	L	H	A	C	K	E	S	-	247	
<i>Cicer_arietinum</i> LOC101508260/1-215	159	A	K	E	C	K	I	V	F	E	Y	G	P	L	I	L	I	N	A	E	K	F	L	E	K	T	A	D	I	C	T	T	L	H	A	C	P	A	S	-	196	
<i>Medicago_truncatula</i> MTR_7g072560/1-242	171	G	S	K	C	K	I	V	L	E	Y	G	P	L	V	F	E	N	A	E	K	F	L	E	K	T	D	I	C	T	A	L	H	A	C	K	E	S	-	208		
<i>Medicago_truncatula</i> MtrunA17_Chr4g001314/1-223	167	K	K	E	C	K	R	M	V	F	E	Y	G	P	L	I	L	V	N	A	E	K	Y	L	K	K	A	D	I	C	T	T	L	H	A	C	P	S	S	-	204	
<i>Medicago_truncatula</i> MTR_029040/1-215	159	K	K	E	C	K	R	M	V	F	E	Y	G	P	L	I	L	V	N	A	E	K	Y	L	K	K	A	D	I	C	T	T	L	H	A	C	P	S	S	-	196	
<i>Trifoliumpratense</i> L195_g026334/1-194	138	T	K	E	C	K	R	M	V	F	E	Y	G	P	L	I	I	I	N	A	E	K	F	L	K	T	N	D	I	C	T	T	I	H	A	C	P	A	S	-	175	
<i>Lotus_japonicus</i> NA 3S9R9/1-216	160	A	K	E	C	K	R	M	V	F	E	Y	G	P	F	V	L	M	N	A	E	K	F	L	K	T	T	G	I	C	T	A	L	H	A	C	P	A	S	-	197	
<i>Phaseolus_vulgaris</i> PHAVU_008G084800g/1-222	173	A	N	K	C	K	R	M	V	F	E	Y	G	P	L	L	V	F	D	N	A	E	K	F	L	E	N	V	D	I	C	T	V	V	H	A	C	P	A	S	-	210
<i>Phaseolus_vulgaris</i> PHAVU_008G0847000g/1-217	164	A	N	K	C	K	R	M	V	L	E	N	G	P	L	V	F	D	S	A	Q	R	F	L	E	S	T	D	I	C	T	A	V	V	Y	Q	I	F	-	201		
<i>Glycine_max</i> GLYMA_09G277100/1-237	164	A	N	K	C	K	R	M	V	L	E	N	G	P	L	V	F	D	N	A	E	K	F	L	E	S	T	D	I	C	T	A	I	Y	A	Q	I	F	-	201		
<i>Glycine_max</i> GLYMA_01G131400/1-216	160	S	K	K	C	K	R	M	V	F	E	Y	G	P	L	I	L	V	K	A	E	K	F	L	E	K	T	A	D	I	C	T	T	L	H	A	C	P	A	S	-	197
<i>Eucalyptus_grandis</i> EUGRSU_Z_K01273/1-227	171	E	D	K	C	K	M	V	F	E	Y	G	P	L	I	L	G	N	E	K	F	L	E	A	A	D	V	C	S	L	L	H	L	C	A	A	T	-	208			
<i>Eucalyptus_grandis</i> EUGRSU_Z_A00687/1-219	164	A	K	K	C	K	M	V	F	E	Y	G	P	L	V	L	A	N	A	E	Q	F	L	E	A	N	D	V	C	T	T	L	H	A	C	K	A	S	-	201		
<i>Gossypium_hirsutum</i> LOC107896756/1-233	188	V	K	K	C	S	L	V	F	E	Y	G	P	L	I	L	A	N	T	E	N	F	L	E	T	T	D	V	C	T	I	L	H	A	C	N	G	A	-	225		
<i>Gossypium_hirsutum</i> LOC107935966/1-227	182	V	K	K	C	K	R	L	V	F	E	Y	G	P	L	I	L	A	N	A	E	H	F	L	E	T	T	D	V	C	T	I	L	H	A	C	D	G	G	-	219	
<i>Gossypium_tomentosum</i> ES332_D10G139500v1/1-233	188	V	K	K	C	S	L	V	F	E	Y	G	P	L	I	L	A	N	T	E	H	F	L	E	T	T	D	V	C	T	I	L	H	A	C	N	G	A	-	225		
<i>Gossypium_tomentosum</i> ES332_A																																										

<i>Amborella trichopoda</i> AMTR_s00007p00225690/1-214	213	CY	214
<i>Amborella trichopoda</i> AMTR_s00062p00198130/1-320	317	-----MILL	320
<i>Cinnamomum micranthum_f_kanehirae</i> CKAN_01065200/1-278	270	-----HQTGKATDV	278
<i>Cinnamomum micranthum_f_kanehirae</i> CKAN_00757300/1-212	197	GIEAGSNSPLLDITSA	212
<i>Anthurium amnicola</i> Psapl_1_1-288	280	RNQTDHMLNVSMESQSRRHPPSTIAQLYLI	288
<i>Anthurium amnicola</i> Psapl_1_2/1-273	255	PPPCPPSPPEKYQFMAKV	273
<i>Zostera marina</i> ZOSMA_381G00120/1-242	240	-----SDI	242
<i>Zostera marina</i> ZOSMA_56G01350/1-232	222	IAAKSSSQAI	232
<i>Dendrobium catenatum</i> MA16_Dca011512/1-222	204	QAESTIGGAALPGSSIHD	222
<i>Dendrobium catenatum</i> MA16_Dca020165/1-215	202	KATEVEQ-----QFLSASA	215
<i>Oryza sativa_subsp_indica</i> Osl_34843/1-245	229	--KRKAFNLF SARKLV RDA	245
<i>Oryza sativa_subsp_indica</i> Osl_19500/1-223	209	----IIPARAGDLGALSAA	223
<i>Oryza sativa_subsp_japonica</i> Osl_12g0112200/1-245	229	--KRKAFNLF SARKLV RDA	245
<i>Oryza sativa_subsp_japonica</i> Osl_05g0334400/1-223	209	----IIPARAGDLGALSAA	223
<i>Brachypodium distachyon</i> BRADL_4g25580v3/1-245	229	--KQSTARSSFEGLRSDA	245
<i>Brachypodium distachyon</i> BRADL_2g04110v3/1-235	222	----IQATTETVLSLAAL	235
<i>Hordeum vulgare_subsp_vulgare</i> NA F2DBE9/1-246	230	--KKRAVGSFFDGGLRSDA	246
<i>Hordeum vulgare_subsp_vulgare</i> NA A0A287KA24/1-238	225	----VQATTETIPLSATL	238
<i>Triticum aestivum</i> NA A0A386MMT5/1-246	230	--KTRAGGSFFDGLRSDA	246
<i>Triticum aestivum</i> NA A0A386FHG9/1-233	223	----MQATIPLSAAL	233
<i>Sorghum bicolor</i> SORBL_3008G032600/1-247	231	--KQKTFSSVLQGLLSDA	247
<i>Sorghum bicolor</i> SORBL_3003G055700/1-227	214	----TQASMETMPLSATL	227
<i>Zea mays</i> Zeammb73_Zm00001d042734/1-240	220	--SRKTFSSMLKGLWSDAWLEG	240
<i>Zea mays</i> Zeammb73_Zm00001d039719/1-229	216	----TQASMEAMPLSAML	229
<i>Aquilegia coerulea</i> AQUUCO_00400489v1/1-223	205	PDGGEVATSSLEKSLVADA	223
<i>Spinacia oleracea</i> SOVF_050110/1-231	213	NTGKQQQSVGEIRIEMVTSS	231
<i>Helianthus annuus</i> HannXRG_Chr10g0286291/1-181	170	-----GPIEEASLVSDN	181
<i>Cynara cardunculus_var_scolymus</i> Cord_003008/1-231	220	-----ASIEEASKISDN	231
<i>Lactuca sativa</i> LSAT_9X38061/1-229	221	-----EQVPLISDN	229
<i>Coffea canephora</i> GSCOC_700023234001/1-294	276	APTAAVLSSSTSETSLRAAS	294
<i>Nicotiana tabacum</i> LOC107812754/1-245	227	ADKELQASPKMQASLHAS	245
<i>Nicotiana tabacum</i> LOC107792809/1-238	221	--VGKEEVLPMKQTSMHAS	238
<i>Solanum tuberosum</i> 102602502/1-242	225	--VDKEQASPMKQTSLHAS	242
<i>Solanum lycopersicum</i> NA A0A3Q7I0Q/1-238	222	--VDKEQASRKQTSLHAS	238
<i>Cicer arietinum</i> LOC101491522/1-279	248	TVHSKSLLELLSLYHGHTSSRMLHLLKIALY	279
<i>Cicer arietinum</i> LOC101508260/1-215	197	IVISQEATIMEEIPMLSDS	215
<i>Medicago truncatula</i> MTR_7g072560/1-242	209	TVVLEKSFSLDLSIFYGNNIIFIRMVQLLKIALF	242
<i>Medicago truncatula</i> MtrunA17_Chr4g001314/1-223	205	TIVSQEATVTEETALFSDS	223
<i>Medicago truncatula</i> MTR_029040/1-215	197	TIVSQEATVTEETALFSDS	215
<i>Trifolium pratense</i> L195_g026334/1-194	176	SIVSQKTTINEEIPMLSDS	194
<i>Lotus japonicus</i> NA J3S9R9/1-216	198	TAVSQEASIMGEIPLSDS	216
<i>Phaseolus vulgaris</i> PHAVU_008G084800g/1-222	211	-----EVASEQALLSDS	222
<i>Phaseolus vulgaris</i> PHAVU_008G0847000g/1-217	202	NSGWPTSLSLRLLLWKQ	217
<i>Glycine max</i> GLYMA_09G27100/1-237	202	NSGWPNANLSFRFLWKQRSNRFVMYLLQKIAIYITTQ	237
<i>Glycine max</i> GLYMA_01G131400/1-216	198	TAVSNKEASIMEVPLISDS	216
<i>Eucalyptus grandis</i> EUGRSUZ_K01273/1-227	209	EIKSEESVPTKEMPLSDS	227
<i>Eucalyptus grandis</i> EUGRSUZ_A00687/1-219	202	--STITDVEVLETSSVVASS	219
<i>Gossypium hirsutum</i> LOC107896756/1-233	226	-----KQTLVADS	233
<i>Gossypium hirsutum</i> LOC107935966/1-227	220	-----KQESVADS	227
<i>Gossypium tomentosum</i> ES332_D10G139500v1/1-233	226	-----KQTLVADS	233
<i>Gossypium tomentosum</i> ES332_A02G005200v1/1-227	220	-----KQESVADS	227
<i>Theobroma cacao</i> TCM_019744/1-228	221	-----KQASVADS	228
<i>Brassica rapa_subsp_peekinensis</i> NA MHDBN0/1-215	204	-----QTVLRPLGLADS	215
<i>Brassica rapa_subsp_peekinensis</i> NA MCGRM9/1-214	203	-----KTVLTQPGTADS	214
<i>Brassica oleracea_var_oleracea</i> NA A0A0D3DSC3/1-229	218	-----KTVLTQPGTADS	229
<i>Brassica oleracea_var_oleracea</i> NA A0A0D3D313/1-216	205	-----QTVLRPLPGSADS	216
<i>Arabidopsis lyrata_subsp_lyrata</i> ARALYDRAFT_486888/1-220	205	----ATHHGYIPTVEALADS	220
<i>Arabidopsis lyrata_subsp_lyrata</i> ARALYDRAFT_666001/1-213	202	-----KSVLRQPELADS	213
<i>Arabidopsis thaliana</i> At5g01800/1-217	202	ATHRDYVPA--VESLADS	217
<i>Arabidopsis thaliana</i> At3g51730/1-213	202	-----KSVLRQPELADS	213
<i>Rosa chinensis</i> RchiOBHm_Chr3g0460961/1-229	211	TVSTMEASSLDIGSMRADS	229
<i>Prunus persica</i> PRUPE_6G290000/1-253	222	VASTEESASPVTVTVLSDSKSRQQRDTGMMEE	253
<i>Malus domestica</i> DVH24_036312/1-296	283	KASITEH-----ISNLKAS	296
<i>Populus trichocarpa</i> POPTR_016G133400/1-242	228	----EDSMEQASAVLKADS	242
<i>Populus trichocarpa</i> POPTR_006G107300/1-242	228	KDSGEQAS--TMLTADS	242
<i>Cucumis sativus</i> Csa_4G331080/1-233	215	PLGDNAVSSVGTVPPLADA	233
<i>Cucumis melo_var_makuwa</i> E5676_scaffold127G001120/1-249	231	PLGDNAVSSVGTVPPLADA	249
<i>Cucumis melo_var_makuwa</i> E6C27_scaffold1166G00310/1-233	215	PLGDNAVSSVGTVPPLADA	233

Conservation

-----000000000010-----

Quality



Consensus

-----LADs-----
TAGSEEASSVEEQPLLADS+SR+++LQ+IAIY++TQ

Occupancy



Figure 3-03. Sequence alignment of plant PSAPLIPs from several selected angiosperms.

One or two representative sequences were selected from each species. Alignment was conducted with Clustal MUSCLE and image was produced by JalView. Color method was Taylor with a conservation level 85%. Annotation was calculated automatically. Conserved cysteines were highlighted in yellow.

Subcellular localization of *Arabidopsis* PSAPLIPs

Sequence alignment of *Arabidopsis* AtPSAPLIP1 and AtPSAPLIP2 with human prosaposin showed alignments with human saposin B, saposin C and saposin D. In human saposins, there are two glycosylation sites in saposin A (N80 and N101), and one glycosylation sites in saposin B (N215), saposin C (N332) and saposin D (N426). The glycosylation sites in saposin B, saposin C and saposin D are in the same position in alignments. No post-translational modifications are predicted for *Arabidopsis* PSAPLIPs in Uniprot. However, sequence alignment results provide putative glycosylation information in *Arabidopsis* PSAPLIPs. In AtPSAPLIP1, the N57 in the first SapB-like domain is aligned with human saposin B glycosylation site (Figure 3-04A), therefore it is speculated that this site might be glycosylated in the plants. The in the second SapB-like domain, P143 is aligned with the corresponding glycosylation site in human saposin C (Figure 3-04A), and therefore this site. This suggests that this site is unlikely to be glycosylated in the plants. And in total there's only one putative site in AtPSAPLIP1 might be glycosylated in the cells. In AtPSAPLIP2, the first aligned site in

the first SapB-like domain is Y corresponding to the human saposin B glycosylation site and the site in the second SapB-like domain is P (Figure 3-04A). This suggests AtPSAPLIP2 is not glycosylated in the plant cells. Both genes were cloned and overexpressed using the 35S promoter in *Arabidopsis*. Total protein was extracted and digested with Endo-Hf to release the N-linked glycosylation from the proteins. The results showed that AtPSAPLIP1 was glycosylated and AtPSAPLIP2 was not glycosylated (Figure 3-04B). Together, these results support the sequence and structural similarity between *Arabidopsis* PSAPLIPs and human prosaposin. As a further expectation, they may share a similar molecular mechanism in lipid interactions.

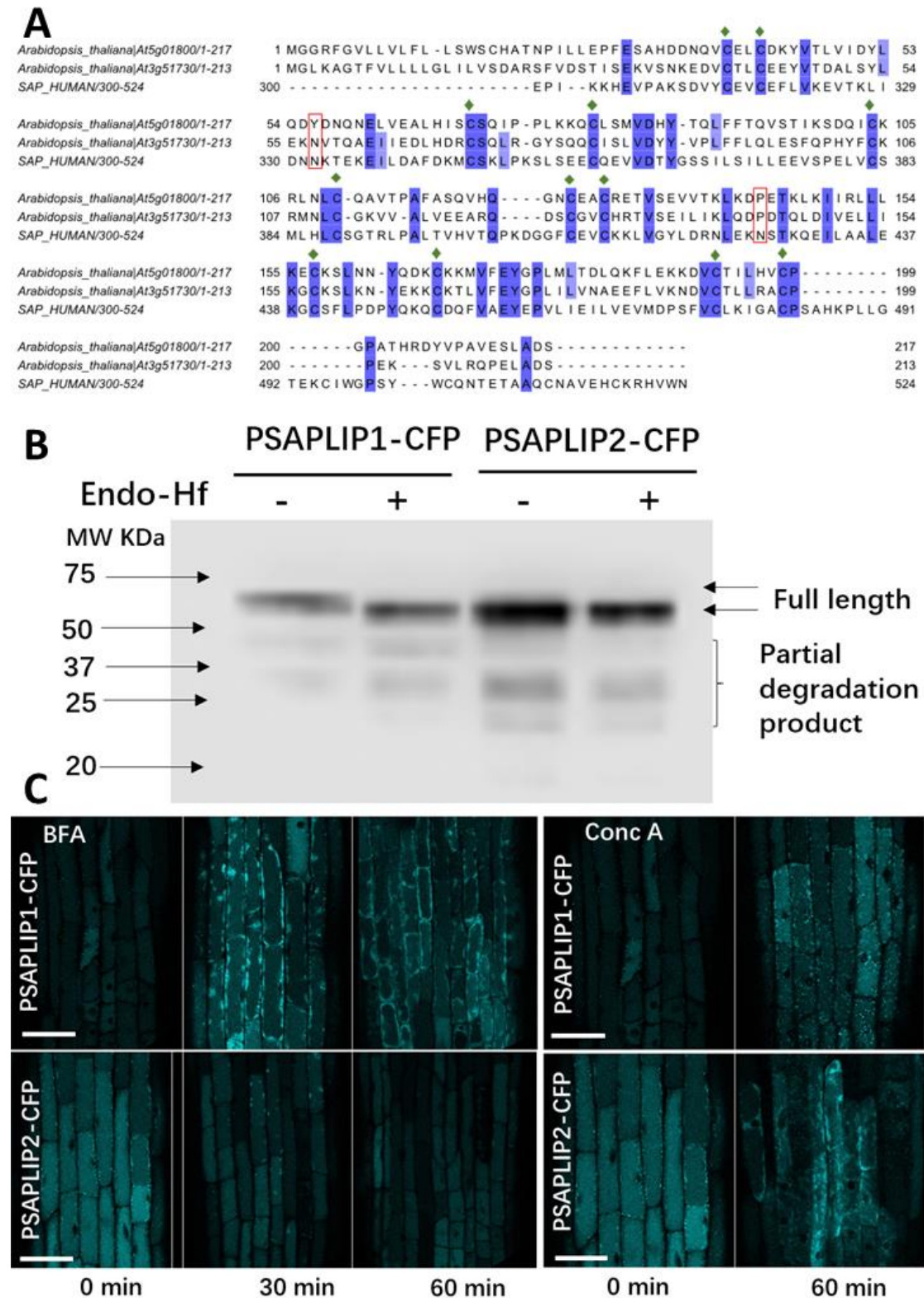


Figure 3-04. Glycosylation and subcellular localization of AtPSAPLIP1 and AtPSAPLIP2.

(A) Sequence alignment between human prosaposin and *Arabidopsis* PSAPLIPs. Red boxes mark the positions of N-glycosylation in human saposins and the corresponding

sites in *Arabidopsis* PSAPLIPs. Green diamonds mark the conserved cysteines. (B) Western blot of AtPSAPLIP1-CFP and AtPSAPLIP2-CFP with and without Endo Hf digestion. Blot was probed with anti-HA. (C) Confocal laser microscopy images of AtPSAPLIP1-CFP and AtPSAPLIP2-CFP. 7 days after germination seedlings were treated with 10 μ M brefeldin A (BFA) or 100nM concanamycin A (conc A). Bar=100 μ m.

Among other species, this putative glycosylation site corresponds to the same glycosylation site as human saposin B. However, this site can also be K, E, H or Y (Figure 3-03; Figure S09), and this suggests that this site is not conserved. This may reflect that glycosylation is not likely to be essential for saposin activities in plants, and the glycosylation may be a trait inherited from ancestors. In *Chlamydomonas reinhardtii*, the corresponding sites are T and E, which also suggests that these two PSAPLIPs may not be glycosylated. The second SapB-like domain corresponding glycosylation site in AtSAPLIP2 is also P, and this site is relatively conserved across plant species. This suggest that plant PSAPLIPs may function in a way different from human saposins.

Since human prosaposins are processed into four mature saposins, it is possible that plant PSAPLIPs may also be like mammalian counterparts. However, whole protein extraction and Western blotting analysis of AtPSAPLIP1 and AtPSAPLIP2 overexpression showed that the amount of full-length protein (58kDa) was much greater than predicted processed single domain (46kDa) (Figure 3-03B). This result suggests that *Arabidopsis* PSAPLIPs might not be processed into single saposin-like

proteins. A “self-dimer” or dimerization with another molecule are likely to occur in the cell. This would function differently from human saposins.

Both AtPSAPLIP1 and AtPSAPLIP2 were targeted to the vacuole (Figure 3-04C). To explore the trafficking pathway of PSAPLIP1 and PSAPLIP2, brefeldin A (BFA) treatment, a fungal inhibitor which blocks trafficking between endoplasmic reticulum (ER) and Golgi complex, were applied. Accumulated signals of PSAPLIP1 appeared after 30 minutes, while PSAPLIP2 did not accumulate after 30 minutes or one hour (Figure 3-04C). These results suggest that PSAPLIP1 passes Golgi body while PSAPLIP2 does not. To test whether PSAPLIP1 and PSAPLIP2 traffic to the vacuole, concanamycin A (concanamycin A) treatment, which inhibits the vacuolar type H-ATPase and further inhibits fusion with vacuoles, were applied. Both proteins were affected and accumulated outside the vacuole after 30 minutes treatment (Figure 3-04C). This result indicates that both proteins traffic to the vacuoles.

Expression pattern of *AtPSAPLIP1* and *AtPSAPLIP2*

To find the possible roles of PSAPLIPs in *Arabidopsis* growth and development, the promoters were cloned and fused with the GUS reporter to elucidate the expression pattern of these two genes. Both genes were highly expressed in floral organs (Figure 3-05 and Figure 3-06).

In floral organs, *AtPSAPLIP1* was primarily expressed in inflorescences, pedicels, receptacles, sepals and the mature pollen, with weak expression in carpels, filaments

and petals (Figure 3-05A, B, C). No expression was detected in stigma or ovules (Figure 3-05D, E). Expression in petals and filaments showed increasing signals with developmental stages (Figure 3-05B). No other expressions were found in anthers except the mature pollen (Figure 3-05C). Signals in germinated pollens were detected on stigmas (Figure 3-05D). *AtPSAPLIP1* expression was previously shown to be upregulated during leaf senescence (Gepstein et al., 2003). However, the promoter GUS results showed that expression was higher in young leaves, while in senescent leaves the staining was weaker (Figure 3-05F). Expression was also detected in the roots (Figure S14).

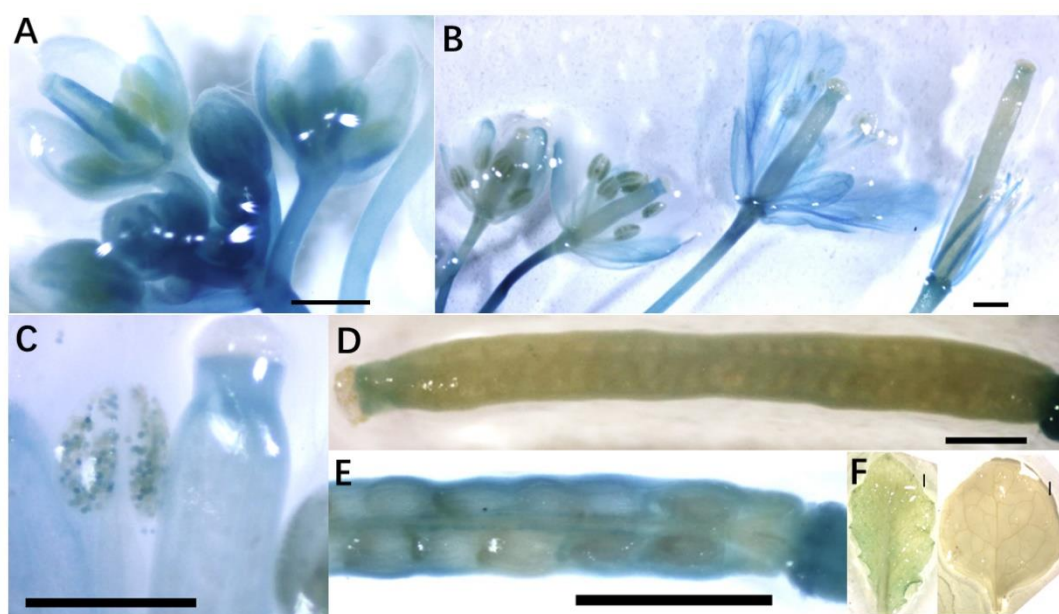


Figure 3-05. *AtPSAPLIP1* promoter GUS staining. (A) Inflorescence and young flower buds. (B) Flowers. From left to right, young to old flowers, with increasing staining in petals and filaments. (C) Mature pollen. (D) 2DAP silique. (E) Developing seeds. (F) Emerging leaf (left) and senescent first true leaf (right). DAP: days after pollination. Bar=1mm.

AtPSAPLIP2 was primarily expressed in inflorescences, pedicels, receptacles, petals, anthers, carpels and ovules, with weak expression in sepals and filaments (Figure 3-06). Expression of *AtSAPLIP2* in petals and anthers showed an interesting pattern: expression was high in anthers and low in petals in younger stages (Figure 3-06B), and decreased in anthers till around flower stage 9 and increased in petals with developmental stage starting from flower stage 8 (Figure 3-06B). Expression in the pollen was not detected (Figure 3-06C). Expression in ovules was present before fertilization (Figure 3-06C). In developing seeds, the expression was detected in young siliques, especially in integuments (Figure 3-06D, E, F, G). The signal decreased rapidly, and approximately 10 days after pollination, no signals were detected in seeds (Figure 3-06H). Developmental stage of embryos approximately matches linear to early mature stages around 10 days after pollination. From Seedgenenetwork.net, *PSAPLIP2* expression is moderate in chalazal seed coat and general seed coat, low in chalazal endosperm at linear stage. Expression is low in embryos before linear stage and not detected in mature embryos. From the GUS staining (Figure 3-06F), the chalaza was stained darker. In vegetative tissues, root tips and leaf veins were also stained (Figure S14).

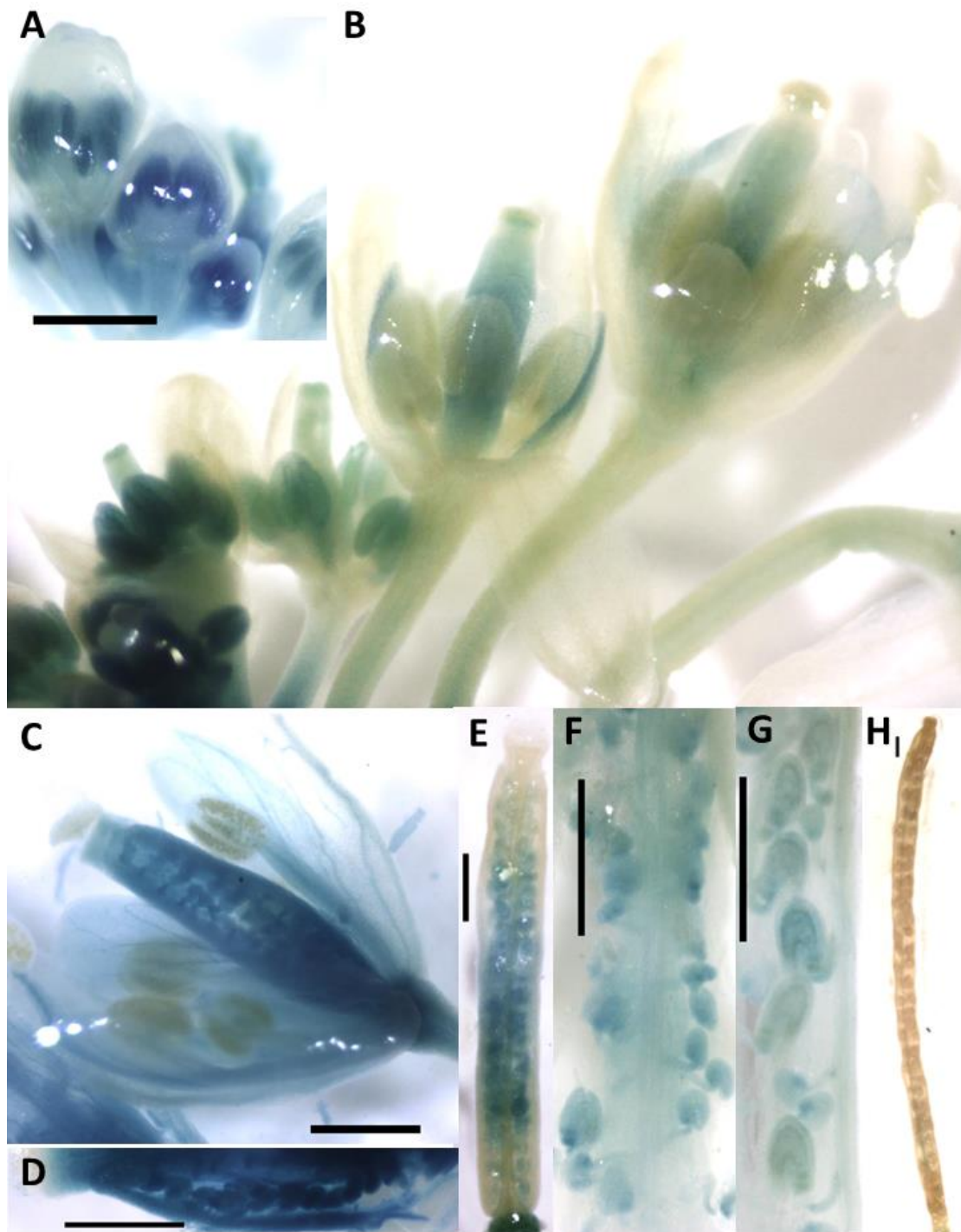


Figure 3-06. *AtSAPLIP2* promoter:: GUS staining. (A) Inflorescence and young flower buds. (B) Flowers. From left to right, young to old flowers, with the changing staining in anthers and petals with different development stages. (C) Flowers at stage 15. Carpel was opened to show the unfertilized ovules. (D) Fertilized ovules. (E) 2DAP

silique. (F) Developing seeds and the integument. (G) 5 DAP silique and the inter integument and endosperms. (H) 10DAP silique. DAP: days after pollination. Bar=1mm.

It appears that one of the most important function of *AtPSAPLIP1* is in pollen maturation, while the function of *AtPSAPLIP2* is in anther development, most likely in pollen formation. *AtPSAPLIP2* is also involved in regulation of early seed development. The involvement in reproductive processes may explain why PSAPLIPs are ubiquitous across the plant kingdom, and the gene copies did not expand since reproductive regulation is a complicated and delicate.

Discussion

PSAPLIPs are ubiquitous in the plant kingdoms

The plant specific insert (PSI) in plant aspartic proteases was the only saposin-like protein found in the plants. The PSAPLIPs in plants remained uncharacterized for a long time. Although the *PSAPLIP* family is not a large gene family, it is ubiquitous in the plant kingdom. This suggests its important role in plant growth and development. However, the function of PSAPLIPs in plant remains unclear.

To elucidate the functions of plant PSAPLIPs, phylogenetic studies were conducted. By searching in protein database Uniprot, around 340 protein sequences were identified as plant PSAPLIPs. While PSAPLIPs containing three SapB-like domains are prevalent from green algae to gymnosperms, the major type of PSAPLIPs in

angiosperms contains only two SapB-like domains. This type first shown in gymnosperms, and this supports the single origin of angiosperms from gymnosperms that containing two SapB-like domains form of PSAPLIPs. In most plant species, there are one to four members in this family in the genome.

Similar to animal PSAPLIPs, plant PSAPLIP also show diverse protein sequences. This was reflected in the phylogenetic tree: the relationship between branches was not always consistent to the phylogenetic relationship between different plant groups. However, the secondary structures of saposin-like domains in plants are highly similar to the mammalian saposins. This was reflected by the comparison between predicted structure of SapB-like domains from *Arabidopsis* PSAPLIPs and mammalian saposin-like protein structure. This highly similar secondary structure may help understanding the molecular function of PSAPLIPs in plant cells. Plant PSAPLIPs may also function as a membrane interactor, similar to human saposins, either by promoting membrane disruption or interaction with other proteins on membranes.

Plant PSAPLIPs are similar to human prosaposin

Some basic features were concluded from the results in *Arabidopsis* PSAPLIPs *AtPSAPLIP1* and *AtPSAPLIP2*. Similar to human saposins in lysosomes, *AtPSAPLIP1* and *AtPSAPLIP2* were localized to vacuoles. The trafficking of both proteins was sensitive to concanamycin A, which supports that they are targeted to the vacuoles. The difference between these two proteins is PSAPLIP1 was sensitive to BFA treatment while PSAPLIP2 was not. This suggests that *AtPSAPLIP1* is transported to *trans*-Golgi

network first before trafficking to vacuoles, and AtPSAPLIP2 is directly trafficked to vacuoles. Unconventional trafficking directly towards vacuoles from ER has been previously reported, and it is dependent on post-translational modification, glycosylation, of the PSI in aspartic proteases (Vieira et al., 2019). Experimental results showed that the AtPSAPLIP1 was glycosylated and AtPSAPLIP2 was not. This may explain the difference between PSAPLIP1 and PSAPLIP2 in trafficking routes. However, how glycosylation affects the trafficking route choice, and whether glycosylation affects saposin-like protein activity are still unclear. It seems that the glycosylation does not significantly affect the saposin activity (Rossman et al., 2008). In the working model of human saposins, glycosylation may help hide the hydrophobic cavity, not it does not appear to be necessary for interaction with lipids (Rossmann et al., 2008). In plant PSAPLIPs, the putative glycosylation site is not conserved either across different species. This also supports the hypothesis that glycosylation seems beneficial but not essential to saposin-like protein activity.

The primary form of mature *Arabidopsis* PSAPLIPs proteins was found to contain two SapB-like domains. The single saposin-like protein versions were much less than the full-length form. This reflects plant PSAPLIPs function in a different way from human prosaposin. Human prosaposin is processed into individual saposins, while in insect cells, di-saposins are the major products of processed prosaposins (Leonova et al., 1996). This indicates that prosaposins are not necessarily processed into individual saposin-like domains to function, and di-saposins are also able to function in the cell.

In plants, di-saposins may be the primary form as the functional unit in the cell. It may form a 'self-dimer' by interaction between the two SapB-like domains or interact with another PSAPLIP molecule in the cell.

PSAPLIPs are likely involved in regulating plant reproductive processes

Human saposins are co-factors for lipid-degradation enzymes (Kishimoto et al., 1992; Schuette et al., 2001). Little is known about degradation of sphingolipids in plants. Overexpression of *PSAPLIPs* in *Arabidopsis* did not affect plant growth and development (Figure S10-S13). One explanation is that PSAPLIPs are not involved in sphingolipid metabolism in plants. However, PSAPLIPs are not the lipid degradation enzymes themselves, and overexpression is not likely to disturb the metabolic levels without additional hydrolases. The proteomic studies on PSAPLIP interacting proteins may help in identify the sphingolipid hydrolases in the plants.

The expression pattern of *AtPSAPLIP1* and *AtPSAPLIP2* suggest that they are important in reproduction processes. The differential expression indicates functional differentiation between *AtPSAPLIP1* and *AtPSAPLIP2*. Animal PSAPLIP biological function is usually identified with the corresponding diseases and it is hard to infer the biological functions in the plants by sequences or structural features. The interaction between PSAPLIPs and other target proteins may help to explore the biological functions of plant PSAPLIPs. Expression and interacting databases provide clues for SAPLIP interacting partners. In the BioGrid database, *AtPSAPLIP1* is annotated as interacting with *AtRACK1A(Receptor for Activated C Kinase 1 A)*. RACK1A is a member

of the tryptophan-aspartate repeat (WD-repeat) family of proteins, function in shuttling proteins around the cell, anchoring proteins at particular locations and in stabilizing protein activity (Adams et al., 2011). RACK1A was reported involved in ABA signaling and may be required for production of ribosomes complex (Guo et al., 2011). This suggests that PSAPLIP1 may function in interaction between the RACK1A and the membrane system.

The expression pattern of *AtPSAPLIP1* also suggests a role in pollen maturation. *AtPSAPLIP2* is annotated as interacting with SYNTAXIN-23 (SYP23), WAVY GROWTH 2 (WAV2) and EXCESS MICROSPOROCTES1 (EMS1). WAV2 is primarily expressed in the roots. It encodes a protein belonging to the BUD EMERGENCE 46 family of proteins with a transmembrane domain at the N terminus and an α/β -hydrolase domain at the C terminus (Mochizuki et al., 2005). This may be one of the lipid hydrolase candidates. The SNARE protein SYP21, SYP22 and SYP23 all localize on vacuolar membrane (Shirakawa et al., 2010). They may be the interactor for facilitating target protein interactions with the vacuolar membrane. *EMS1* is expressed in tapetum, inner integument and chalaza. This expression pattern overlaps with *AtPSAPLIP2* expression pattern (Figure 3-06). EMS1 is a leucin-rich receptor-like kinase which is localized on plasma membrane. It can interact with the ligand TAPETUM DEVELOPMENT 1 (TPD1) in regulation tapetum development (Huang et al., 2016). It is likely that *AtPSAPLIP2* interacts with endocytosed EMS1 and transports it to vacuoles to terminate signal transduction. The primary biological function of *AtPSAPLIP1* and *AtPSAPLIP2* may be

in tapetum development and pollen maturation. The knockout mutant analysis is essential to further explore the role of PSAPLIPs in plants.

Conclusion

Plant PSAPLIPs are a small gene family in the plant kingdom. It shows similarity to the human prosaposin, in terms of protein structures, post-translational modification and subcellular localization. Plant PSAPLIPs may share a similar molecular mechanism of lipid interaction to human saposins. The expression pattern suggests their important role in flower development, but the exact role remains unresolved.

Till now no T-DNA insertional lines are available for both genes. Generation mutants with CRISPR is one of the alternative choices. Second, the proteomic studies help identifying the biological pathways that AtPSAPLIPs are involved in, such as lipid metabolism or male gametophyte development signaling. Third, structural studies and lipid interaction assays *in vitro* will explore how plant PSAPLIPs interact with lipids and elucidate the biophysical and biochemical properties of these proteins. It is likely that *Arabidopsis* PSAPLIPs form a 'self-dimer' structure with the two SapB-like domains. Whether there is oligomerization is also of interest. Trafficking of PSAPLIPs in the cell is also important, since PSAPLIP2 appear to adopt an unconventional trafficking pathway in the cell. Whether the trafficking pathway affects the function of PSAPLIPs remains unclear. Some of the potential PSAPLIPs interactors such as EMS1, are targeted to the plasma membranes. Where the PSAPLIPs and the target proteins meet

in the trafficking pathways need further explore. Other directions also deserve attention, such as plant defense response. PSI from aspartic proteases is shown to have anti-bacteria activity *in vitro* and enhance the plants resistance to bacteria pathogens (Muñoz et al., 2010; Frey et al., 2018). However, the independent function of PSI from the aspartic protease *in vivo* is not reported yet. Plant PSAPLIPs may also show anti-bacteria activity and function in plant defense responses. If PSAPLIPs are involved in anti-bacterial activity *in vivo*, then it could be considered a good candidate for enhancing resistance to a broad spectrum of bacteria pathogens.

The study of PSAPLIPs from other plant species will also broaden the understanding of their biological functions in plants. For those species lacking floral structures, whether PSAPLIPs take part in reproductive process needs to be explored. This might contribute to our understanding that how reproduction system changes during evolution.

In summary, this dissertation first characterized some features of plant prosaposins-like proteins and provides new insights on how plants regulate reproductive process.

Materials and Methods

Primary and Secondary Structure Prediction

Hydropathy plot was drawn in ExPASy with Kyte and Doolittle method. Window size was 9 with the linear weight variation model. Structure prediction was conducted

in Phyre2. Each SapB-like domain was predicted separately. Predicted structure of AtPSAPLIP1 and AtPSAPLIP2. Final image was visualized with EzMol.

Sequence Alignment

PSAPLIPs protein sequences were selected in EggNOG and Uniprot. In EggNOG, sequences were found by screening orthologs with *AT3G51730*. 167 sequences from 67 species were outputs. In Uniprot, sequences were screened by the keyword saposin. Only sequences in Viridiplantae were chosen for further screening. The sequences annotated as fragments were removed. Aspartic proteases were removed as well. For those without a gene ID, if sequence similarity was above 95%, the longer one was kept. If the annotated SapB-like domains length was below 50 amino acid residues, the corresponding sequences were also removed.

Some sequences were removed because they belong to the neucleophosmin family. This may result from incorrect auto-prediction in Uniprot. The remaining sequences were considered valid PSAPLIP proteins in plants and were selected for further analysis.

Alignment was conducted in MegaX with Clustal MUSCLE method. The parameters were as following: gap open -2.9, gap extend 0, hydrophobicity multiplier 1.2, max memory in MB 2048, max iterations 16, cluster method UPGMA, cluster method UPGMA, min diag length 24. Some manual adjustments were applied for gap positions for better alignments.

To better search for conserved positions, the sequences that only contain one SapB-like domain were removed. Sequences in green algae, liverworts, mosses and gymnosperms were aligned separately due to their variable number of copies of SapB-like domains. Human prosaposin and *Arabidopsis thaliana* PSAPLIPs were chosen as the outlier in this case.

Images were processed with JalView. Color was added by Taylor method with conservation level 85%. Annotation was calculated automatically.

Phylogenetic tree construction

Phylogenetic tree of plant PSAPLIPs were constructed in MegaX with maximum likelihood method. Phylogeny test was bootstrap method, with 2000 bootstrap replications. Substitutions type was amino acid with WAG model. Rates among sites were uniform. All sites were considered. ML heuristic method was nearest neighbor interchange method. No branch swap filter. Number of threads was 3.

Preparation of Transgenic Plants

AtPSAPLIP1 and *AtPSAPLIP2* coding sequences (CDS) were cloned from *Arabidopsis* flower cDNA. The fragments were incorporate into pDONR/Zeo by BP reaction according to the manufacturer's instructions. The entry clones were confirmed by sequencing and the primers for cloning and sequencing are listed in Table S01. The entry clones were then incorporated into pEarleyGate102 (35S promoter, with CFP and HA tag on C-terminus) by LR reaction. The expression constructs were

confirmed by sequencing and the sequencing primers are listed in Table S01. The expression constructs were transformed into agrobacterium strain GV3101, and floral dipped into *Arabidopsis* flowers. The positive seedlings were screened by hygromycin selection and confirmed by confocal microscopy.

Promoter GUS reporter lines were constructed in a similar way. For *AtPSAPLIP1*, the promoter includes 3 prime UTR of previous gene and 5 prime UTR of *PSAPLIP1*, with a total length 451bp. For *AtPSAPLIP2*, the promoter includes the 5 prime UTR with a total length 1.5kb. The fragments were also inserted into pDONR/Zeo by BP reaction and incorporated into pGWB3 by LR reaction. Both entry clones and expression constructs were confirmed by sequencing. The cloning and sequencing primers are listed in Table S01. Transgenic plants were screened by hygromycin selection. Details on methods of molecular clonings are described in appendix D.

Plant Materials and chemical treatment

All the *Arabidopsis thaliana* plants were in Columbia-0 (Col-0) ecotype genetic background. For germination on solid media, *Arabidopsis* seeds were surface sterilized by soaking in 20% bleach (containing sodium hypochlorite) for 15 minutes with agitation. Seeds were then rinsed three to five times in sterile water. Seeds were sowed on 1/4 Murashige and Skoog (MS) medium (RPI Corp.) media containing 0.5% sucrose with 0.8% agar. Seeds were stratified at 4°C for 2 days in the dark and then placed in growth chamber at 22°C, with 24 hr continuous white light at 100 $\mu\text{mol m}^{-2} \text{s}^{-1}$. For

chemical treatment, seedlings were first grown for 4 days on regular 1/4MS media, then transferred to new media containing the corresponding chemicals. The working concentrations of chemicals used in this research were: 10 μ M brefeldin A (BFA); 100nM concanamycin A (conc A); 4 μ M propidium iodide (PI); 5 μ g/ml fluorescein diacetate (FDA); 2 μ M abscisic acid (ABA); 75mM sodium chloride (NaCl). Low nitrogen media was prepared by adding 10 μ M potassium nitrate (KNO₃) in 1/4MS without nitrogen (MS w/o nitrogen) media; sufficient nitrogen media was prepared by adding 5mM KNO₃ in 1/4MS w/o nitrogen media. Ethanol or DMSO was added as the control, with the concentration same as the corresponding chemicals. Adult plants were grown in growth chamber at 22°C with 16 hr light and 8 hr darkness cycles.

Protein extraction

300mg *Arabidopsis* seedling tissues were ground with a grind stick in Eppendorf tubes with liquid nitrogen. The ground tissues were resuspended in 300 μ L protein extraction buffer (50 mM sodium citrate, pH 5.5; 5% SDS (w/v); 0.01% BSA (w/v); 150 mM NaCl; 2% (v/v) β -mercaptoethanol and 1 μ L of protease inhibitor cocktail (Genesee Scientific). The mixture was incubated for 60 minutes at 100° C. Samples were centrifuged at 4° C, 14,500g for 30 minutes and the supernatant was collected. The samples were stored in -80° C if not used immediately.

Glycosylation test

Glycosylation was detected by Endo Hf (New England BioLabs) digestion

according to manufacturer's instruction. Briefly, 17 μ L extracted protein sample was added with 2 μ L 10xGlycoBuffer 3, 1 μ L Endo Hf. Samples were incubated at 37°C for 1 hour. Then the sample was for SDS-PAGE and Western blot.

SDS-PAGE

Total proteins were separated by SDS polyacrylamide gel electrophoresis (SDS-PAGE). 10 μ L samples were prepared by adding 2 μ L of 6X SDS (sodium dodecyl sulfate) loading buffer (1.2g SDS, 0.01% bromophenol blue, 4.7ml glycerol, 1.2ml Tris 0.5M pH=6.8, 2.1ml water). Samples were loaded onto 12% polyacrylamide 0.75mm 10-well or 15-well gel (Bio-Rad®). Precision Plus Protein Dual Color Standards (Bio Rad) was used as size markers. Electrophoresis was carried out in 1X Running Buffer (3g of Tris base, 14.4g of glycine, and 1g of SDS in 1000 ml water) at 120V for approximately 4 hours or until the dye front reached the front of the gel.

Western blot

For immunoblotting, proteins were transferred to polyvinylidene difluoride (PVDF) membrane in Tris-glycine-methanol transfer buffer (2.9g glycine, 5.8g Tris, 0.37g SDS 100mL methanol, 900mL water) at 120V for 80 minutes at 4°C and then rinsed briefly in 1xPBS. Membranes were blocked overnight at 4°C in blocking buffer (5% non-fat milk in 1xPBS with 0.02% Tween20) or 1.5 hours at room temperature. The membrane was rinsed gently with washing buffer (1% non-fat milk in 1x PBS with 0.02% Tween20) for three times each for 15 minutes. The membrane then was incubated with primary

antibody (anti-HA) in blocking buffer overnight at 4°C or 1.5 hours at room temperature. The membrane was rinsed with washing buffer for three times and each for 15 minutes. Then the membrane was incubated with secondary antibody (anti-rabbit digoxigenin) at room temperature for 1.5 hours. The membrane was rinsed with washing buffer for three times and each for 15 minutes. Proteins were visualized using a SuperSignal West Femto Kit (Thermo Scientific). Images were taken by C-DiGit Blot Scanner (LI-COR).

Microscopy

Confocal microscopy was carried out using a Zeiss LSM 710 Confocal laser scanning microscope (Carl Zeiss, Germany) with Axio Imager 2. Pixel dwell time was 0.01 ms. The master gain was always set to less than 893, with a digital gain of 1.5. For RFP/mCherry acquisition: 594 nm (5%) excitation and 588-696 nm emission. For YFP acquisition: 514 nm (5%) excitation and 519-560 nm emission. For GFP: 488 nm (5%) excitation and 493-598 nm emission. For CFP: 458 nm (5%) excitation and 453-580 nm emission. For PI: 543 nm (5%) excitation and 583-718 nm emission. For FDA: 488 nm (5%) excitation and 493-583 nm emission. Quantification of fluorescence intensity was analyzed using ZEN Lite 2012. Briefly, representative images from 10 nuclei in each of the 5

Seven to 10 anthers from stage 13-14 were selected for imaging. All images were processed with ZEN Lite 2012 (Zeiss) and ImageJ.

Histochemistry

For GUS staining detection, plant tissues were fixed in cold 90% acetone for 30 minutes, then washed twice in GUS buffer before staining. Samples were infiltrated with GUS buffer under vacuum for 10 minutes, then incubated at 37°C for 48 hours. Tissue was cleared in 70% ethanol overnight and repeated several times until the tissue becomes clean and clear. The sample was mounted on microscope slides for visualization.

Chapter 4: Conclusion and Perspective

Plants have two types of proteins contain saposin B-like domains: aspartic proteases with the plant specific insert (PSI) and prosaposin-like proteins (PSAPLIPs). In this dissertation, three main questions were addressed. What are the biological functions of these aspartic proteases in plants? What is the role of the saposin-like domain (plant specific insert) in those processes? What are the biological functions of prosaposin-like proteins in plants? Using molecular genetic analysis, I conclude that the typical aspartic proteases function in bulk proteolytic activity such as seed storage protein processing/degradation and programmed cell death (PCD). The plant specific insert guides the protease towards the vacuole and perhaps also facilitates membrane disturbance. There is no evidence supporting that this PSI could function independently from the protease. From phylogenetic analysis and studies in *Arabidopsis AtPSAPLIPs*, I conclude that PSAPLIPs are important in reproductive processes especially in male gametophyte development.

First, I began phenotypic studies in the least studied *Arabidopsis* aspartic protease *ASPA2*. *ASPA2* is expressed throughout the plant, which is similar with the reported expression pattern of *ASPA1*. Single loss-of-function *aspa2* mutant showed delayed seed maturation. As the delay in maturation is subtle, I suspected that there was redundancy in three *ASPAs* in *Arabidopsis*. Then I tested phenotype of *aspa1-2 aspa2-1 aspa3-3* triple mutant (*ASPA1* is knock-down, *ASPA2* and *ASPA3* are knock-out alleles). Triple mutant seeds showed delayed germination in terms of germination rate and

seed storage proteins degradation. The fusion of small vacuoles to form the central vacuole was also delayed in the mutant cotyledons. This result suggests that ASPAs involved in not only proteolytic activity but also membrane disturbance.

To explore *ASPAs* function in other tissues, I compared root growth in the wild type, triple mutant and *ASPAs* overexpression plants. Root architecture was different in response to nitrogen supply in that the triple mutant root showed more lateral roots and primary root growth was relatively insensitive to nitrogen levels compared to wild type. Further analysis suggested that the altered root architecture may result from tracheary element (TE) maturation in xylem tissues. The triple mutant showed slight delay TE maturation and the *ASPA2* overexpression showed slightly earlier maturation. Together with the expression pattern of *ASPA3*, this indicates that *ASPAs* may regulate the rate of programmed cell death in *Arabidopsis*.

Then I monitored PCD in lateral root cap with propidium iodide (PI) staining or propidium iodide/ fluorescein diacetate (PI/FDA) double staining. The distance between stained nuclei and root tip was not different between triple mutant and wild type. This suggests that onset of PCD was not delayed in the mutant. But the distance between the distal nucleus and root tip was longer in the mutant which indicated a longer execution time of PCD. This was confirmed by time-course imaging with PI staining. To test the mechanism of *ASPAs* in PCD, PI/FDA double staining was applied to monitor the time from cytosolic pH drop to PI signal appearance in nucleus. This time was longer in the triple mutant and shorter in overexpression plants. This

indicates that membrane permeability increased more slowly in the mutants and faster in the overexpression plants. This reflects the role of ASPAs in the rates of membrane permeability regulation during PCD.

The ASPA promoters were cloned for constructs with a HISTONE 2A 10 reporter tag constructs to detect transcriptional responses to stress signals. *ASPA1* expression remained constant while *ASPA2* expression was upregulated by low nitrogen and ABA treatment but downregulated by salt stress. This result indicated that *ASPA1* is more like a housekeeping gene and *ASPA2* is more responsive to different signals for fine tuning plant growth and development to stress signals. *ASPA3* expression was confined to tissues that undergo PCD, and most likely to function in protein degradation and nitrogen recycling to fine-tune the rate of PCD.

The independent function of PSI was not detected in this dissertation. No phenotype was found in triple mature mutant or catalytic inactive *ASPA2* overexpression plants. For future studies, it would be good to create a knock-out triple mutant for analysis. Since the *ASPA1* allele is knock-down, the remaining *ASPA1* activity may compensate of the missing ones and thus the phenotype of *aspa1-2/2-1/3-3* may be weak. Second, one would complement the phenotype of knock-out plants. If there is a triple knock-out mutant, the catalytic inactive version *ASPA2* D107A would also be used to determine if the PSI itself might rescue the phenotype. The reason may be that the phenotypes of the knock-down mutant was weak, and it was not powerful to test the difference when *ASPA2* D107A was overexpressed in plants. Third, one would

screen for the substrates of APSAs to discover if membrane proteins were targets. The PSI allows the protease to associate with membranes, and this close interaction brings the protease and substrates together and makes it possible for membrane protein degradation.

Through sequence screening and alignment, I found that prosaposin-like proteins (PSAPLIPs) are ubiquitous throughout plant kingdom. This family did not disappear, nor did it expand in evolution either. In most species, there are 1-4 genes in the genome. In angiosperms, there is an N-terminal signal peptide and two saposin B (SapB)-like domains. In gymnosperms, liverworts, mosses and green algae, PSAPLIPs contain three SapB-like domains. Plant PSAPLIPs show low sequence similarity but high similarity in secondary structure of SapB-like domains. This structural similarity was indicated by glycosylation analysis of *Arabidopsis AtPSAPLIP1* and *AtPSAPLIP2*, and *AtPSAPLIP1* was glycosylated and *AtPSAPLIP2* was not. Both *AtPSAPLIP1* and *AtPSAPLIP2* were targeted in vacuoles. However, trafficking of *AtPSAPLIP1* was sensitive to BFA while *AtPSAPLIP2* was not. The differences in glycosylation and response to trafficking inhibitors indicate different trafficking routes to the vacuole. Although they are trafficked in different routes, both are in the vacuole, and this indicates that PSAPLIPs function in facilitating degradation of specific proteins.

Then the promoter GUS reporter constructs were created for expression analysis. *AtPSAPLIP1* was primarily expressed in inflorescence, especially in sepals, carpels and mature pollen. *AtPSAPLIP2* was expressed in inflorescence too, but primarily in young

anthers, petals and ovules. These results suggest differential functions of PSAPLIPs in *Arabidopsis*. Since both are expressed in stamens, the results indicate the role in male gametophyte development. This may explain why this family is widely spread in the plant kingdom but did expand during evolution. Future studies include the molecular genetic analysis using loss-of-function mutants. So far no T-DNA insertional mutants are available for those two genes, CRISPR would be employed to create mutants. The potential male sterility phenotype would be the focus of the biological observation. Preliminary data suggest that there is a possible *AtPSAPLIP2* mutant which shows male sterility phenotype. Confirmation of the mutation via sequencing is needed. Second is the proteomic study to find protein-protein interactions. One possible target for *AtPSAPLIP2* is *EXCESS MICROSPOROCTES1 (EMS1)* due to annotation in plant protein-protein interaction data base BioGrid and the overlapping expression patterns in flowers. The third, in plants where floral structures do not exist, such as liverworts and mosses, the functions of PSAPLIPs need to be explored, and this may also provide information about the reason why there is one SapB-like domain missing in angiosperms.

Summary

This dissertation was divided into two parts. Part one is the first *in vivo* study of ASPA biological functions in the plants. Besides demonstrating ASPAs role in both seed development and seed germination *in vivo* for the first time, this is also the first

time showing that ASPAs are involved in programmed cell death in plants. Part two characterized some features of the plant prosaposin-like proteins (PSAPLIPs) for the first time. The potential role in male gametophyte development may contribute to our knowledge of the regulation of plant reproductive processes.

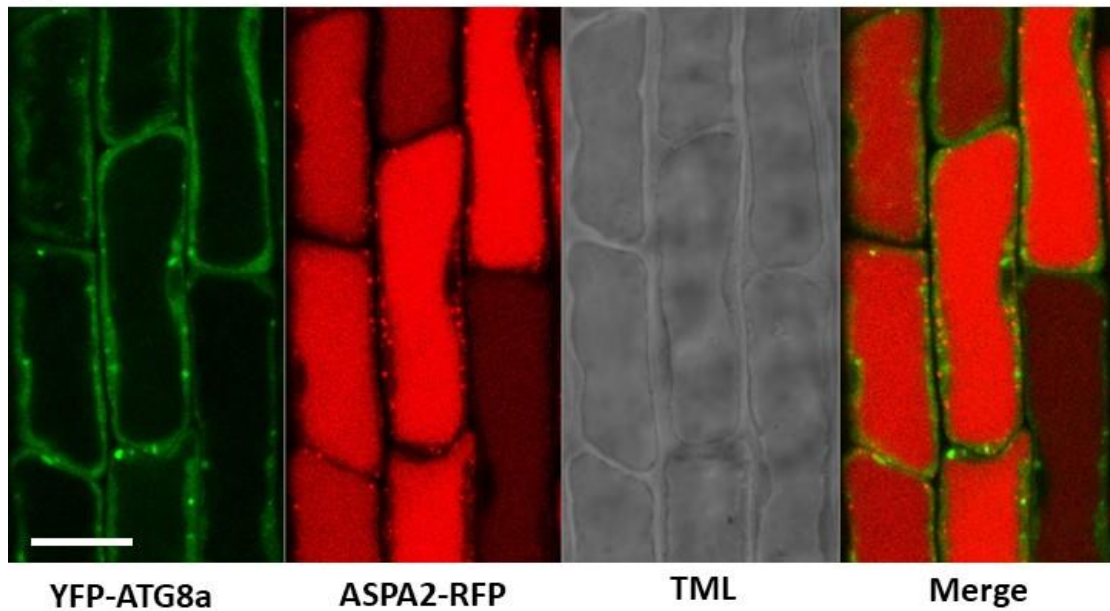
In the first part, three hypotheses were made: ASPA2 was involved in regulation of seed development; ASPAs were involved in regulation of programmed cell death; the plant specific insert (PSI) in ASPAs has an independent biological function *in vivo*. By molecular genetic studies, results showed that ASPAs regulated seed development and seed germination. ASPAs also regulated programmed cell death in lateral root caps, and they are likely to promote membrane permeability in this process. By studying the catalytic inactive form of ASPA2 overexpression plants, the supporting results for an independent role of PSI were not obtained. Further research directions would include proteomic studies on whether ASPAs prefer degradation of membrane targets. Phenotypic studies on the impact of PCD defects in other tissues such as the tapetum, ideally with the knockout triple mutants. Cellular and molecular studies on whether there are other trafficking pathways for ASPAs such as secretion to the extracellular space, and how glycosylation affects ASPAs in the cell. And in other species, whether ASPAs have other biological functions is another important direction.

In the second part, plant PSAPLIPs were characterized for the first time. Some features, such as the distribution in the plant kingdom, the sequence structures, the

predicted secondary structures were described. By molecular genetic studies of *Arabidopsis* PSAPLIP1 and PSAPLIP2, their functions were proposed as interaction with and facilitating target proteins for degradations. They were important in male gametophyte development. Further directions of studies include phenotypic studies on the knockout mutants, elucidating the biological functions in different tissues. Proteomic studies are also helpful for screening the interactors for potential targets and help to understand in which pathways are PSAPLIPs in the cells. The structural studies will increase our understanding on how these proteins function, especially the interaction with lipids. This may help to explore the common and unique features of plant PSAPLIPs compared from animal PSAPLIPs.

Appendix A Supplemental Figures for Chapter 2

A



B

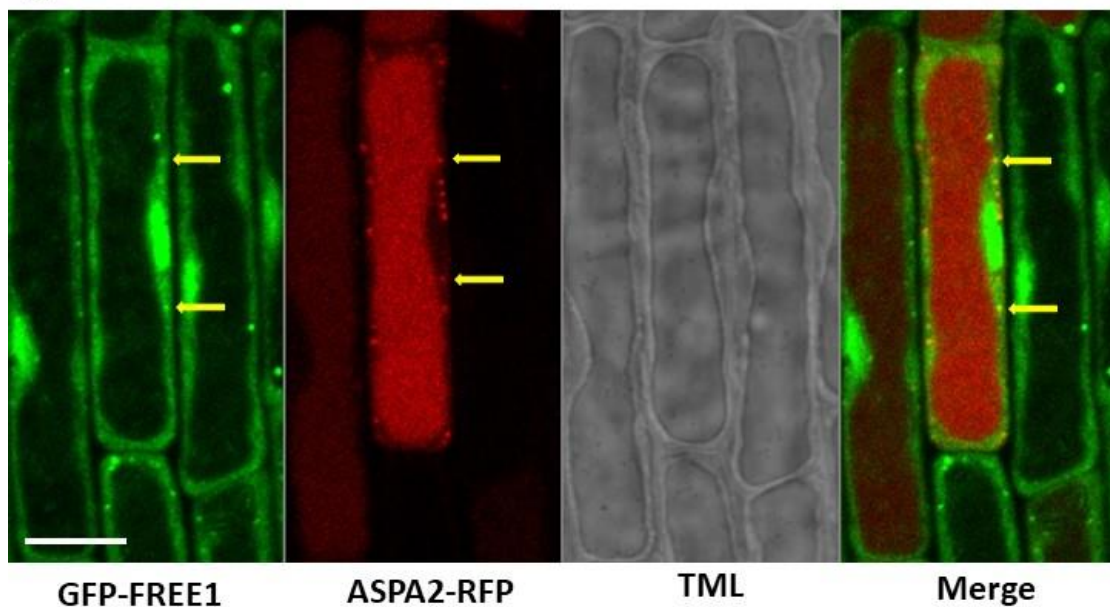


Figure S01. Colocalization between ASPA2 and autophagy marker ATG8a, and endosomal sorting complex required for transport (ESCRT) machinery associated protein FREE1. (A) Colocalization between YFP-ATG8a and ASPA2-RFP. Pearson'

correlation r range: -0.25 – 0.03 (B) Colocalization between GFP-FREE1 and ASPA2-RFP.

Pearson's correlation r range: 0.17 – 0.26. Yellow arrows point to the overlapped

signals. 5 DAG seedlings were treated with 100nM conc A for 1 hour and imaged.

Bar=10 μ m.

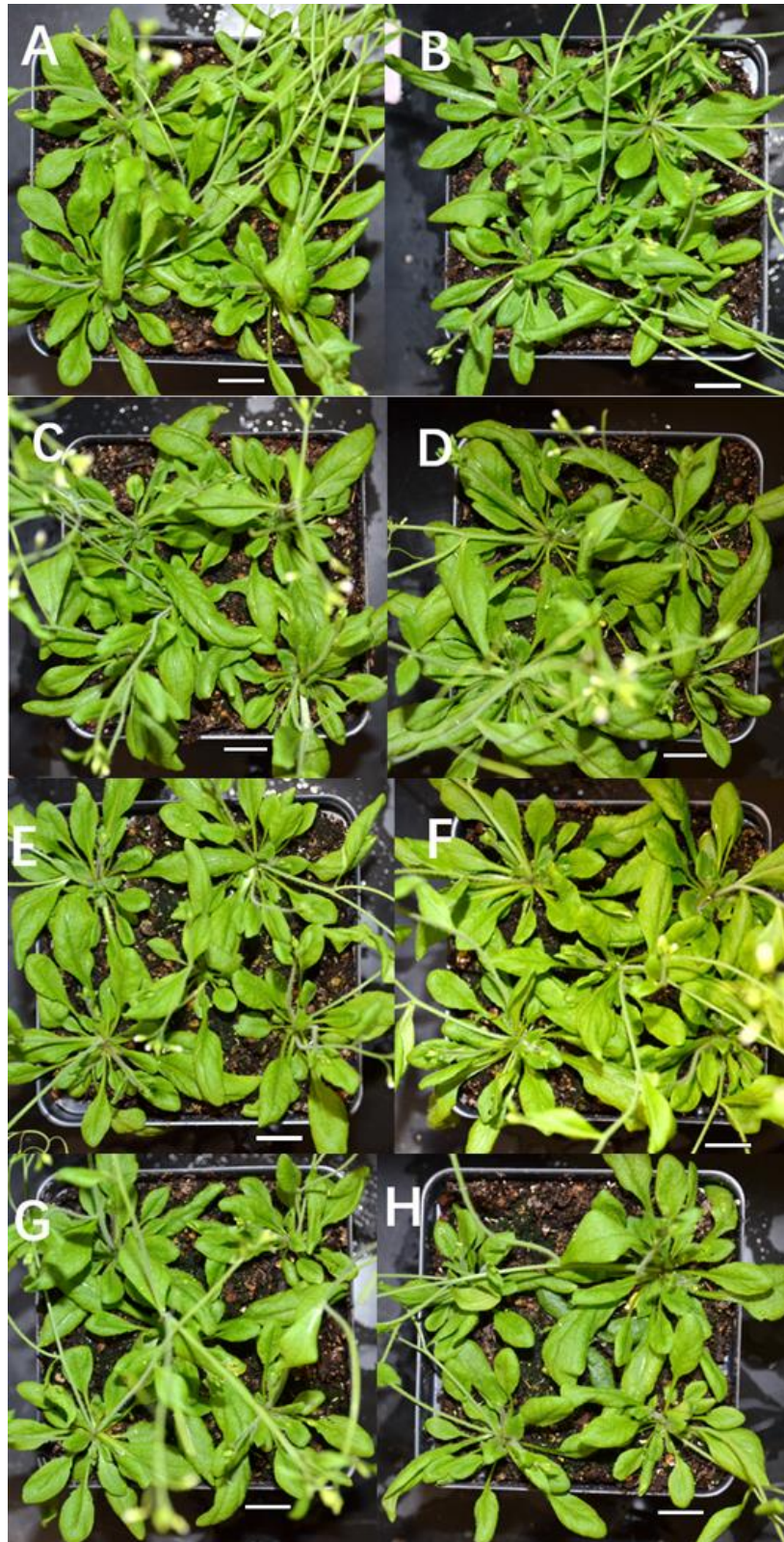


Figure S02. Phenotype of 30 DAG plants of *ASPA* overexpression lines. (A) Col (B) *aspa1-2/2-1/3-3* (C) 35S::*ASPA1*-RFP (D) 35S::*ASPA2*-RFP (E) 35S::*ASPA2* D107A-RFP (F) 35S::*ASPA2* D107A R402Q-RFP (G) 35S::*ASPA2* 321AA-RFP (H) 35S::*ASPA3*-RFP.

Bar=2cm.

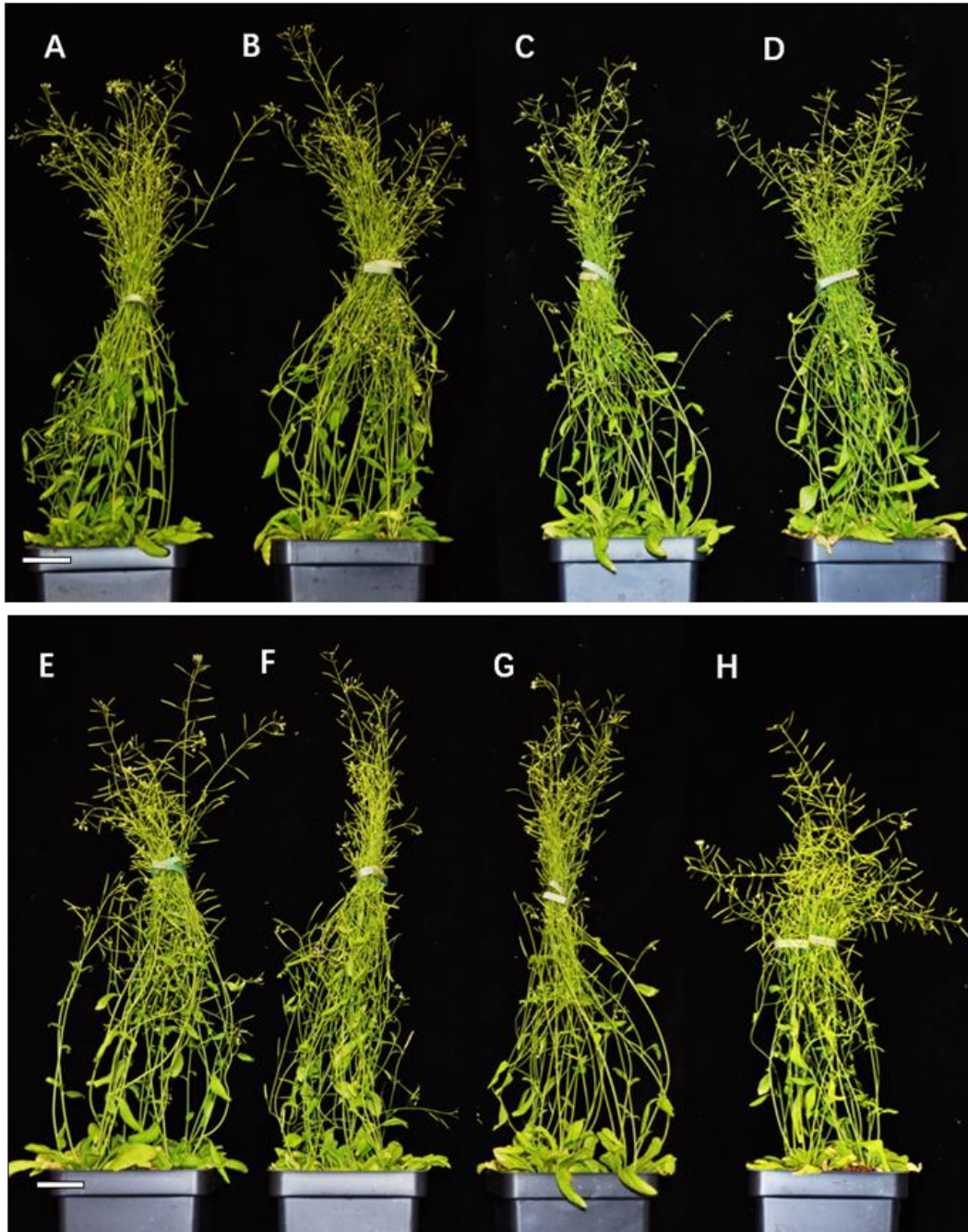


Figure S03. Phenotype of 40 DAG plants of *ASPA* overexpression plants. (A) Col (B) *aspa1-2/2-1/3-3* (C) 35S::*ASPA1*-RFP (D) 35S::*ASPA2321AA*-RFP (E) 35S::*ASPA2*-RFP (F) 35S::*ASPA2*-D107A-RFP (G) 35S::*ASPA2* D107A R402Q-RFP (H) 35S::*ASPA3*-RFP.

Bar=5cm.

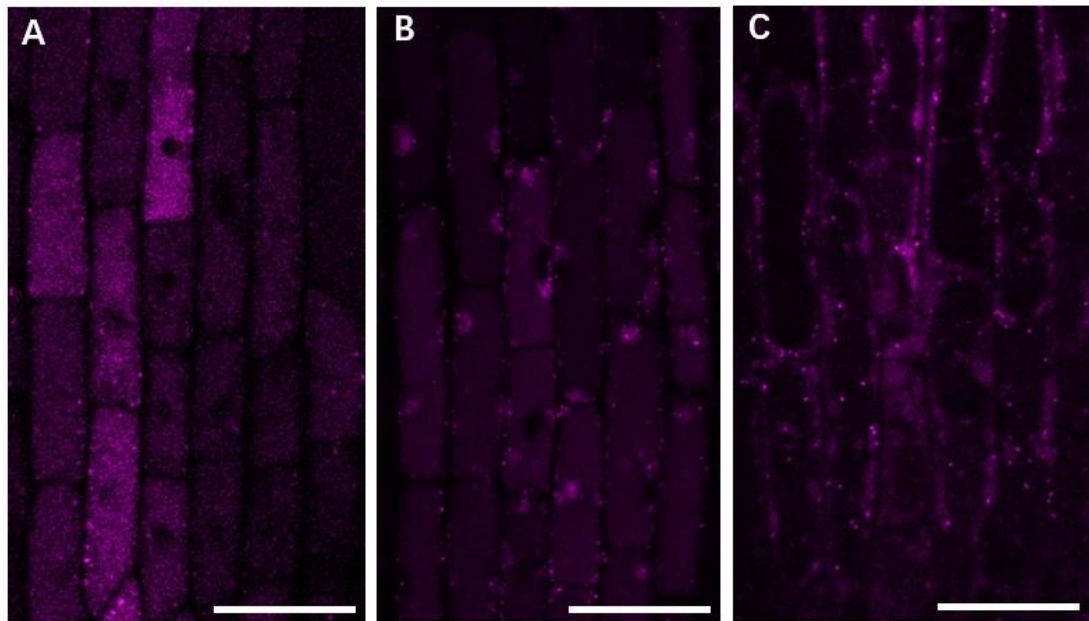


Figure S04. 35S promoter::ASPA2 D107A N404A-CFP (potential glycosylation site mutation) subcellular localization. (A) Control. (B) BFA treatment for 30 min. (C) Conc A treatment for 30min. Bar=20 μ m.

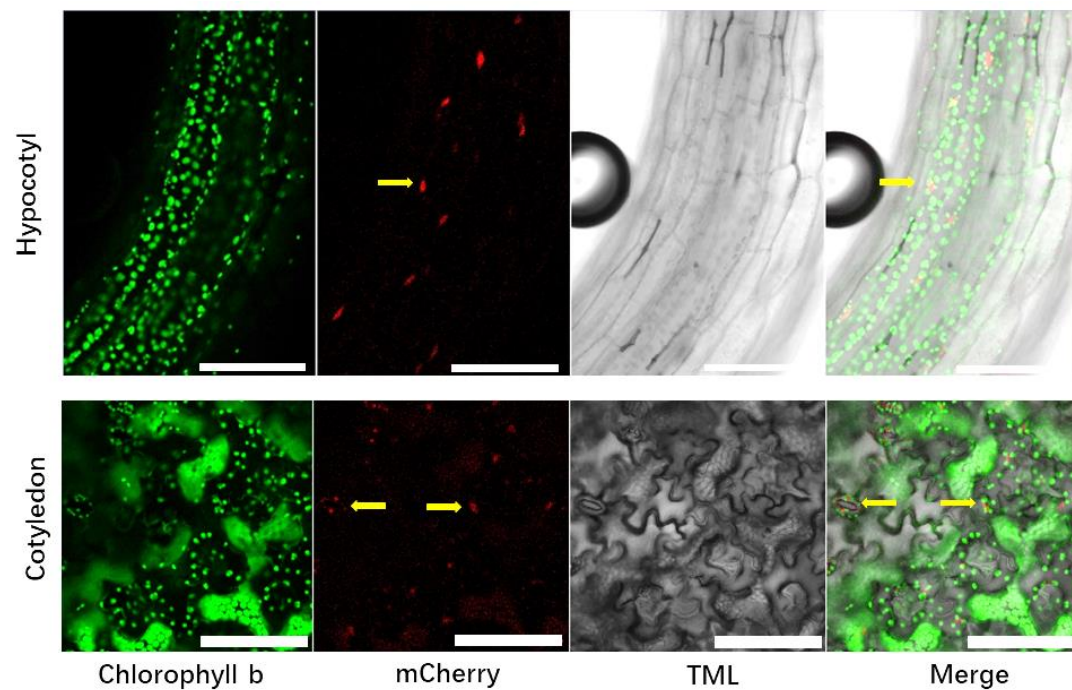


Figure S05. *ASPA1* expression in seedlings. Hypocotyls (top) and cotyledon (bottom).

Yellow arrows point to the signals in nuclei. Bar=200 μ m.

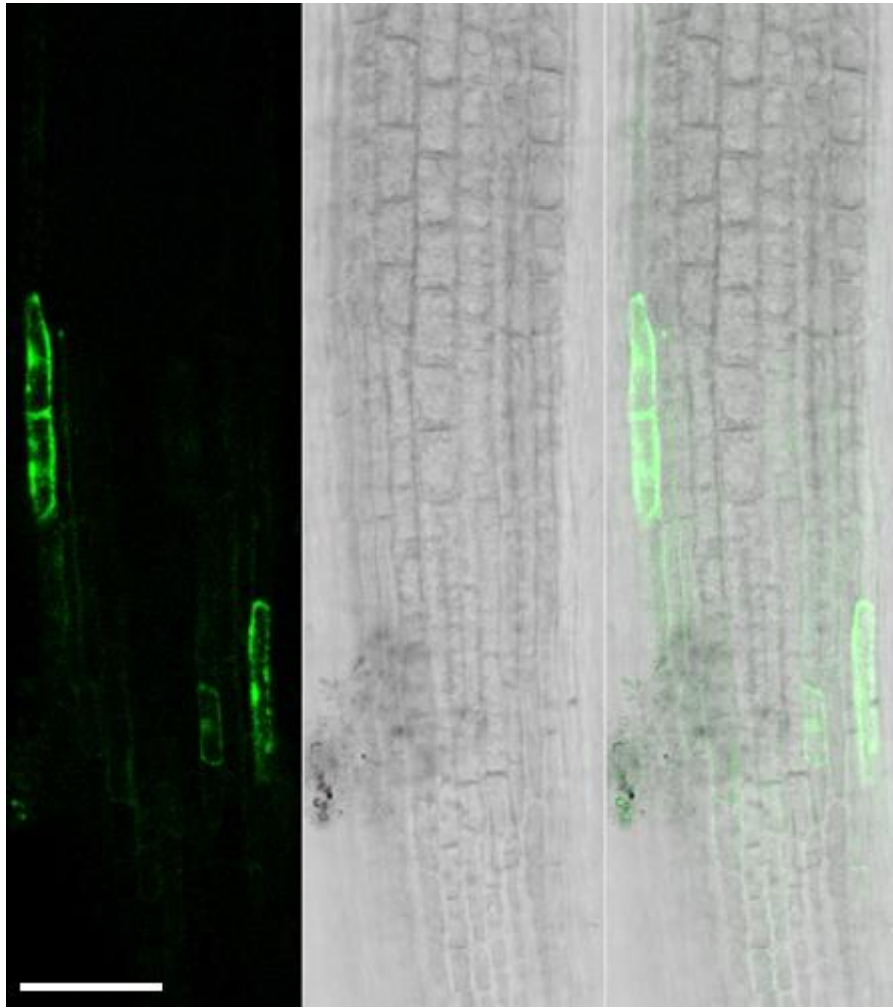


Figure S06. *ASPA3* promoter::YFP expression in lateral root cap. Bar=50 μ m.

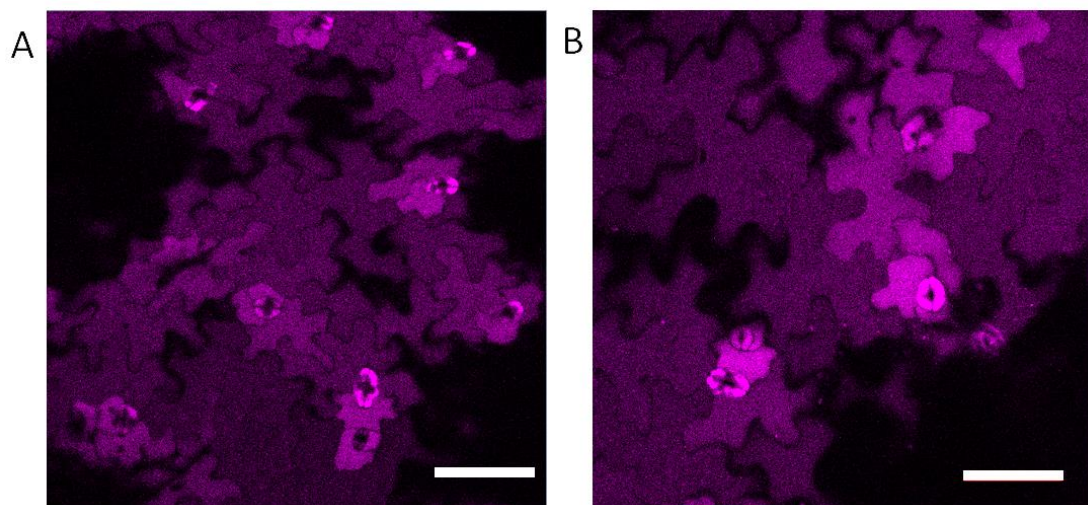
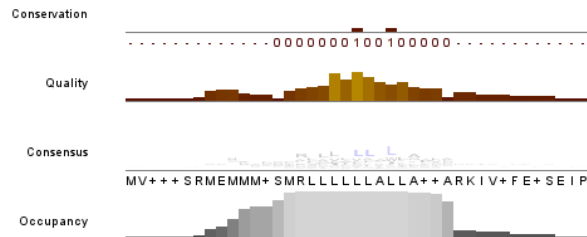


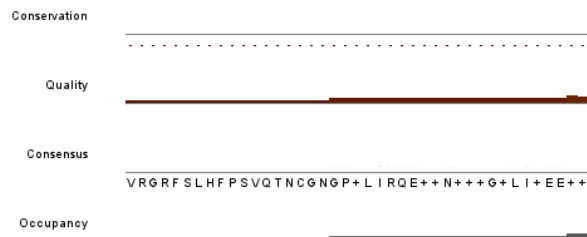
Figure S07. Subcellular localization of ASPA1 and ASPA3 in mature leaves. (A) ASPA1-CFP. (B) ASPA3-CFP. Bar=50 μ m.

Appendix B Supplemental Figures for Chapter 3

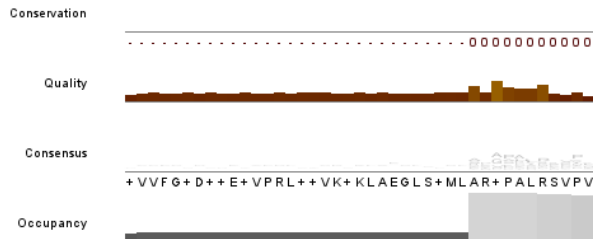
Chloropocon_primus[A3770_07P47130/1-390
Chloropocon_primus[A3770_02p14820/1-272
Chloropocon_primus[A3770_04p29840/1-218
Chloropicon_primus[A3770_04p29830/1-216
Tetrademus_obliquus[BQ4739_LOCUS15021/1-209
Raphidocelis_subcapitata[Rauc_10640/1-361
Monoraphidium_neglectum[MNE_G_12603/1-188
*Coccomyxa_subellipsoidea*_(strain_C-169)[COCSDRAFT_45864/1-332
Chlorella_variabilis[CHLNCDRAFT_58828/1-332
Chlorella_sorokiniana[C2E21_8413/1-327
*Auxenochlorella_protothecoides*_(Chlorella_protothecoides)[APUTEX25_001631/1-331
Tetraselmis_sp.[GSL018]TSPGSL018_26319/1-371
*Micromonas_puilla*_(strain_CCMP1545)[MCPUCDRAFT_47615/1-328
*Micromonas_commoda*_(strain_RCC299/NOUM17/CCMP2709)[MCPUN_62224/1-149
*Micromonas_commoda*_(strain_RCC299/NOUM17/CCMP2709)[MCPUN_105899/1-283
Ostreococcus_tauri[BE221DRAFT_194138/1-242
Bathycoccus_prasinos[Bathy10g00200/1-188
Bathycoccus_prasinos[Bathy07g01820/1-312
Gonium_pectorale[GPECTOR_69g440/1-516
Tetrabaena_socialis[TSOC_008198/1-430
Chlamydomonas_reinhardtii[CHLRE_05g235700v5/1-429
Chlamydomonas_reinhardtii[CHLRE_02g105200v5/1-462
Chlamydomonas_eustigma[CEUSTIGMA_g11715.1/1-345
Klebsormidium_nitens[KFL_001110040/1-305
Chara_braunii[CBR_g3540/1-391
Physcomitrella_patens_subsp._patens[PHYPA_022478/1-333
Physcomitrella_patens_subsp._patens[PHYPA_018982/1-334
Wollemia_nobilis[NA/1-408
Araucaria_cunninghamii[NA/1-406
Picea_sitchensis[NA][A9NUE1/1-430
Picea_sitchensis[NA][A9P228/1-326
Amborella_trichopoda[AMTR_s00007p00225690/1-214
Amborella_trichopoda[AMTR_s00062p00198130/1-320



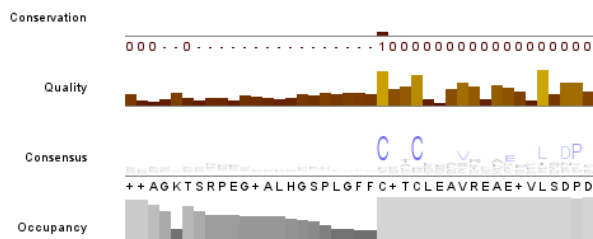
Chloropocon_primus[A3770_07P47130/1-390
Chloropocon_primus[A3770_02p14820/1-272
Chloropocon_primus[A3770_04p29840/1-218
Chloropicon_primus[A3770_04p29830/1-216
Tetrademus_obliquus[BQ4739_LOCUS15021/1-209
Raphidocelis_subcapitata[Rauc_10640/1-361
Monoraphidium_neglectum[MNE_G_12603/1-188
*Coccomyxa_subellipsoidea*_(strain_C-169)[COCSDRAFT_45864/1-332
Chlorella_variabilis[CHLNCDRAFT_58828/1-332
Chlorella_sorokiniana[C2E21_8413/1-327
*Auxenochlorella_protothecoides*_(Chlorella_protothecoides)[APUTEX25_001631/1-331
Tetraselmis_sp.[GSL018]TSPGSL018_26319/1-371
*Micromonas_puilla*_(strain_CCMP1545)[MCPUCDRAFT_47615/1-328
*Micromonas_commoda*_(strain_RCC299/NOUM17/CCMP2709)[MCPUN_62224/1-149
*Micromonas_commoda*_(strain_RCC299/NOUM17/CCMP2709)[MCPUN_105899/1-283
Ostreococcus_tauri[BE221DRAFT_194138/1-242
Bathycoccus_prasinos[Bathy10g00200/1-188
Bathycoccus_prasinos[Bathy07g01820/1-312
Gonium_pectorale[GPECTOR_69g440/1-516
Tetrabaena_socialis[TSOC_008198/1-430
Chlamydomonas_reinhardtii[CHLRE_05g235700v5/1-429
Chlamydomonas_reinhardtii[CHLRE_02g105200v5/1-462
Chlamydomonas_eustigma[CEUSTIGMA_g11715.1/1-345
Klebsormidium_nitens[KFL_001110040/1-305
Chara_braunii[CBR_g3540/1-391
Physcomitrella_patens_subsp._patens[PHYPA_022478/1-333
Physcomitrella_patens_subsp._patens[PHYPA_018982/1-334
Wollemia_nobilis[NA/1-408
Araucaria_cunninghamii[NA/1-406
Picea_sitchensis[NA][A9NUE1/1-430
Picea_sitchensis[NA][A9P228/1-326
Amborella_trichopoda[AMTR_s00007p00225690/1-214
Amborella_trichopoda[AMTR_s00062p00198130/1-320



Chloropocon_primus[A3770_07P47130/1-390
Chloropocon_primus[A3770_02p14820/1-272
Chloropocon_primus[A3770_04p29840/1-218
Chloropicon_primus[A3770_04p29830/1-216
*Tetrademus_obliquus*BQ4739_LOCUS15021/1-209
*Raphidocelis_subcapitata*Rauc_10640/1-361
Monoraphidium_neglectum[MNEG_12603/1-188
Coccomyxa_subellipsoidea(strain_C-169)[COC SUBDRAFT_45864/1-332
*Chlorella_variabilis*CHLNC DRAFT_58828/1-332
Chlorella_sorokiniana[C2E21_8413/1-327
Auxenochlorella_protothecoides(*Chlorella_protothecoides*)[APUTEX25_001631/1-331
*Tetraselmis*_sp.GSL018[TSPGSL018_26319/1-371
Micromonas_pusilla(strain_CCMP1545)[MCPUC DRAFT_47615/1-328
Micromonas_commoda(strain_RCC299/NOUM17/CCMP2709)[MCPUN_62224/1-149
Micromonas_commoda(strain_RCC299/NOUM17/CCMP2709)[MCPUN_105899/1-283
Ostreococcus_tauri[BE221DRAFT_194138/1-242
Bathyococcus_prasinus[Bathy10g00200/1-188
Bathyococcus_prasinus[Bathy07g01820/1-312
Gonium_pectorale[GPECTOR_69g440/1-516
Tetraena_socialis[TSOC_008198/1-430
Chlamydomonas_reinhardtii[CHLRE_05g235700v5/1-429
Chlamydomonas_reinhardtii[CHLRE_02g105200v5/1-462
Chlamydomonas_eustigma[CEUSTIGMA_g11715.t1/1-345
Klebsomidium_nitens[KFL_001110040/1-305
Chara_braunii[CBR_g3540/1-391
Phycomitrella_patens_subsp._patens[PHYPA_022478/1-333
Phycomitrella_patens_subsp._patens[PHYPA_018982/1-334
Wollemia_nobilis[NA/1-408
Araucaria_cunninghamii[NA/1-406
Picea_sitchensis[NA]A9NUE1/1-430
Picea_sitchensis[NA]A9P228/1-326
Amborella_trichopoda[AMTR_s00007p00225690/1-214
Amborella_trichopoda[AMTR_s00062p00198130/1-320

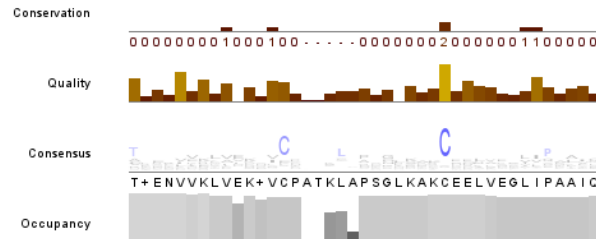


Chloropocon_primus[A3770_07P47130/1-390
Chloropocon_primus[A3770_02p14820/1-272
Chloropocon_primus[A3770_04p29840/1-218
Chloropicon_primus[A3770_04p29830/1-216
*Tetrademus_obliquus*BQ4739_LOCUS15021/1-209
*Raphidocelis_subcapitata*Rauc_10640/1-361
Monoraphidium_neglectum[MNEG_12603/1-188
Coccomyxa_subellipsoidea(strain_C-169)[COC SUBDRAFT_45864/1-332
*Chlorella_variabilis*CHLNC DRAFT_58828/1-332
Chlorella_sorokiniana[C2E21_8413/1-327
Auxenochlorella_protothecoides(*Chlorella_protothecoides*)[APUTEX25_001631/1-331
*Tetraselmis*_sp.GSL018[TSPGSL018_26319/1-371
Micromonas_pusilla(strain_CCMP1545)[MCPUC DRAFT_47615/1-328
Micromonas_commoda(strain_RCC299/NOUM17/CCMP2709)[MCPUN_62224/1-149
Micromonas_commoda(strain_RCC299/NOUM17/CCMP2709)[MCPUN_105899/1-283
Ostreococcus_tauri[BE221DRAFT_194138/1-242
Bathyococcus_prasinus[Bathy10g00200/1-188
Bathyococcus_prasinus[Bathy07g01820/1-312
Gonium_pectorale[GPECTOR_69g440/1-516
Tetraena_socialis[TSOC_008198/1-430
Chlamydomonas_reinhardtii[CHLRE_05g235700v5/1-429
Chlamydomonas_reinhardtii[CHLRE_02g105200v5/1-462
Chlamydomonas_eustigma[CEUSTIGMA_g11715.t1/1-345
Klebsomidium_nitens[KFL_001110040/1-305
Chara_braunii[CBR_g3540/1-391
Phycomitrella_patens_subsp._patens[PHYPA_022478/1-333
Phycomitrella_patens_subsp._patens[PHYPA_018982/1-334
Wollemia_nobilis[NA/1-408
Araucaria_cunninghamii[NA/1-406
Picea_sitchensis[NA]A9NUE1/1-430
Picea_sitchensis[NA]A9P228/1-326
Amborella_trichopoda[AMTR_s00007p00225690/1-214
Amborella_trichopoda[AMTR_s00062p00198130/1-320



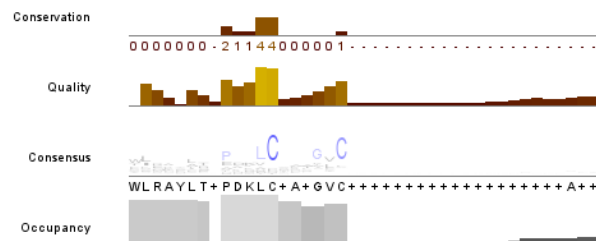
Chloropocon_primus[A3770_07P47130/1-390
Chloropocon_primus[A3770_02p14820/1-272
Chloropocon_primus[A3770_04p29840/1-218
Chloropicon_primus[A3770_04p29830/1-216
Tetrademus_obliquus[BQ4739_LOCUS15021/1-209
Raphidocelis_subcapitata[Rscu_10640/1-361
Monoraphidium_neglectum[MNEG_12603/1-188
*Coccomyxa_subellipsoidea*_(strain_C-169)[COCSUBDRAFT_45864/1-332
Chlorella_variabilis[CHLNCDRAFT_58828/1-332
Chlorella_sorokiniana[C2E21_8413/1-327
*Auxenochlorella_protothecoides*_(Chlorella_protothecoides)[APUTEX25_001631/1-331
*Tetraselmis*_sp.GSL018[TSPGSL018_26319/1-371
*Micromonas_pusilla*_(strain_CCMP1545)[MCPUCDRAFT_47615/1-328
*Micromonas_commoda*_(strain_RCC299/NOUM17/CCMP2709)[MCPUN_62224/1-149
*Micromonas_commoda*_(strain_RCC299/NOUM17/CCMP2709)[MCPUN_105899/1-283
Ostreococcus_taui[BE221DRAFT_194138/1-242
Bathycoccus_prasinus[Bathy10g00200/1-188
Bathycoccus_prasinus[Bathy07g01820/1-312
Gonium_pectorale[GPECTOR_69g440/1-516
Tetrabaena_socialis[TSOC_008198/1-430
Chlamydomonas_reinhardtii[CHLRE_05g235700v5/1-429
Chlamydomonas_reinhardtii[CHLRE_02g105200v5/1-462
Chlamydomonas_eustigma[CEUSTIGMA_g117151/1-345
Klebsormidium_nitens[KFL_001110040/1-305
Chara_braunii[CBR_g3540/1-391
*Physcomitrella_patens*_subsp._patens[PHYPA_022478/1-333
*Physcomitrella_patens*_subsp._patens[PHYPA_018982/1-334
Wollemia_nobilis[NA/1-408
Araucaria_cunninghamii[NA/1-406
Picea_sitchensis[NA|A9NUE1/1-430
Picea_sitchensis[NA|A9P228/1-326
Amborella_trichopoda[AMTR_s00007p00225690/1-214
Amborella_trichopoda[AMTR_s00062p00198130/1-320

73 RPDELDMSKERDGLL--KALQKGMKGA--NGERRRVVPPYL 111
44 TEVPALRITIT--PQKAVVAENQ----- 65
71 TYAAFKKYVEDKVCT--AF--RLVERILCKKVAPKLIPIAYE 108
63 TEEKITKVVNDVCP--KL--PANIAGICASYAPMLLSYAIQ 100
53 VDDTLAKWMVDNICS-----NFDEKEQSDSLVLGLTALVQ 88
47 VEGDVEDWVIGNVCP-----ATGNEKSCADVVTGIIAPALFD 82
46 VEGNVVDWVIDNVCA---AAGDNKQSCADIVNGIAPALLD 82
60 TQSEIIGMILKDVCP--KL--PADAQEAQGGLAPSLIPLGVM 97
62 TQATVEEYIESSVCA--GL--PDQFAQMCTQEVPLVLAQAAD 99
63 TQQTISKYIEAAACA--GL--PDNFKQMKQKEVPVLVASFQ 100
57 TQAYVQELLVGAMCE--GL--PDAFHDTQVQELAAVYRQVVA 94
69 RPLEL--KLGVGAAP--GW--DEGLEGMCKEKRRLLVPPSL 105
71 TQEKFAQYAVKACET-----AGKDKNMCEQLIASGIGSATQ 106
.....
59 TLNAITDKVNE--VCE--TYAGTKHEAMCESLLEMYVQKAI 96
71 FGGDVARAIIARACEEATGG--SESQMGVCVAAGEAGLRFATR 110
56 KFIGPEEVVIESMSN-----VCEVKK-----R 77
54 FERGRKE-----EDERKTVGESILGKTVPPTFT 81
55 ATDFLVDFVEKQICP-----AVGDTAQCHNLAEGLLPTLVQ 90
33 ATEFLVDFVERQICP-----AVGDSVKCHNLAEGLLPTLIQ 68
70 AVAFVVDLF EKQLCP-----ATPDKDECEQLAEAFIPVAMQ 105
56 ATEFLVDLVEKQVCP-----AMGDSAQCHNLAEGLLPTVIQ 91
50 SMDAMVSLVGNNICT--ALAVGKKVDTCRSMSSLLLPAFSR 88
60 NINAVVSLAESQLCL--KLVDPLVSKRELVEEYIPALFQ 98
111 AVDVAVQSFEQFVCG--KLQAEGIREKCKEYIYIPALID 149
69 TPQKVIDKADEFVCH--SL--QPGLKKKCEKMAVEYVPAAIL 106
71 TSMKVMKTADHVLC--KL--QPGLKTKCERMVADYVPAAIL 108
136 TLKNIIEELT--KNLCK--SL--PSNFSAQCEMSQMYIQEIA 172
134 TLKNIIEELT--KSI CK--SL--PSNFSAQCEMSQMYIQETIA 170
156 TLENVAVKLA--KSI CN--EL--PSDLSAKCDEMLGTYIQEVVS 192
156 TLENVAVKLA--KSI CN--EL--PSDLSAKCDEMLGTYIQEVVS 192
39 TFLTITR----- 46
114 FLENIKKCA--GNICS--LLPSNLQGECEESFKSYIEKAVV 150



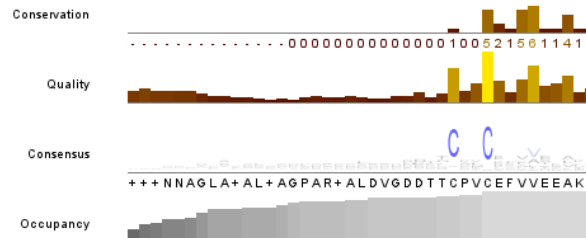
Chloropocon_primus[A3770_07P47130/1-390
Chloropocon_primus[A3770_02p14820/1-272
Chloropocon_primus[A3770_04p29840/1-218
Chloropicon_primus[A3770_04p29830/1-216
Tetrademus_obliquus[BQ4739_LOCUS15021/1-209
Raphidocelis_subcapitata[Rscu_10640/1-361
Monoraphidium_neglectum[MNEG_12603/1-188
*Coccomyxa_subellipsoidea*_(strain_C-169)[COCSUBDRAFT_45864/1-332
Chlorella_variabilis[CHLNCDRAFT_58828/1-332
Chlorella_sorokiniana[C2E21_8413/1-327
*Auxenochlorella_protothecoides*_(Chlorella_protothecoides)[APUTEX25_001631/1-331
*Tetraselmis*_sp.GSL018[TSPGSL018_26319/1-371
*Micromonas_pusilla*_(strain_CCMP1545)[MCPUCDRAFT_47615/1-328
*Micromonas_commoda*_(strain_RCC299/NOUM17/CCMP2709)[MCPUN_62224/1-149
*Micromonas_commoda*_(strain_RCC299/NOUM17/CCMP2709)[MCPUN_105899/1-283
Ostreococcus_taui[BE221DRAFT_194138/1-242
Bathycoccus_prasinus[Bathy10g00200/1-188
Bathycoccus_prasinus[Bathy07g01820/1-312
Gonium_pectorale[GPECTOR_69g440/1-516
Tetrabaena_socialis[TSOC_008198/1-430
Chlamydomonas_reinhardtii[CHLRE_05g235700v5/1-429
Chlamydomonas_reinhardtii[CHLRE_02g105200v5/1-462
Chlamydomonas_eustigma[CEUSTIGMA_g117151/1-345
Klebsormidium_nitens[KFL_001110040/1-305
Chara_braunii[CBR_g3540/1-391
*Physcomitrella_patens*_subsp._patens[PHYPA_022478/1-333
*Physcomitrella_patens*_subsp._patens[PHYPA_018982/1-334
Wollemia_nobilis[NA/1-408
Araucaria_cunninghamii[NA/1-406
Picea_sitchensis[NA|A9NUE1/1-430
Picea_sitchensis[NA|A9P228/1-326
Amborella_trichopoda[AMTR_s00007p00225690/1-214
Amborella_trichopoda[AMTR_s00062p00198130/1-320

112 YRTGEAVAEQDFPSETIEL-----RV 132
66-----PAGLD----- 70
109 WAEKFLE--TGRAC--GAF-- 124
101 TVEHEID--TGKLCDP--LC-- 116
89 WLRVND--PDTLCSGMGVC-- 106
83 WLRLGTD--ADAMCAEVGV-- 100
83 WLRLGTD--AQEMCAEAGVC-- 100
98 YIQSL--ANELCADATLC-- 114
100 LIEKTL--PKDTCAMGIC-- 117
101 SLSEALD--PQGVCGLLGV-- 118
95 TVVAALD--PHDVCTLAGVC-- 112
106 TSGAAAE--RHGLPQDTTVV-- 123
107 FINDNVT--PQSAQDVAGFC-- 134
30-----PQVLC----- 34
97 VINTDFT--PDVLCADAKLCPEEEAPKEEEAPNDGETRAATE 136
111 WIENHPELEEQAQDALDM-- 137
78 FITYNP--PPEMQNA----- 91
82 LEEEDTTFEENT----- 94
91 WFRASAT--PASLCSAGVC-- 108
69 WFRASAT--PASLCSVGV-- 86
106 WLRASET--PASLCAAVGVCGAALLGDPTWDRKHAGNLQQLT 145
92 WLRASET--PASLCSGAGVC-- 109
89 WFKAAAS--PAHLCSAPSAC-- 106
99 IMATEIT--PAKVCGA--VC-- 114
150 VLREDVT--EDKVCGALKLC-- 174
107 ELETLLG--PEKLCYESGV-- 124
109 ELEALLG--PQKLCYESGLC-- 128
173 MMQDYLS--EDKLCISTGLC-- 190
171 MLQDYLS--EDKLCVSTGLC-- 188
193 TLQDYLS--QDKLCIGTGLC-- 210
193 TLQDYLS--QDKLCIGTGLC-- 210
46-----DERVCTY----- 52
151 FLQEYLS--GERLGNSTGLC-- 168



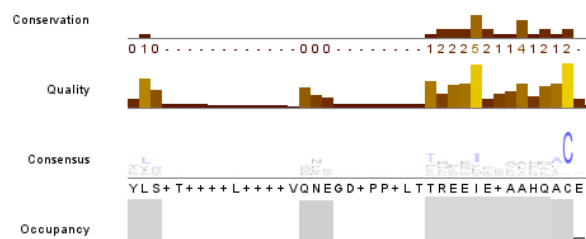
Chloropocon_primus[A3770_07P47130/1-390
Chloropocon_primus[A3770_02p14820/1-272
Chloropocon_primus[A3770_04p29840/1-218
Chloropocon_primus[A3770_04p29830/1-216
Tetrademus_obliquus[BQ4739_LOCUS1502/1-1-209
Raphidocelis_subcapitata[Rauc_10640/1-361
Monoraphidium_neglectum[MNEG_12603/1-188
*Coccomyxa_subellipsoidea*_(strain_C-169)[COGSUBDRAFT_45864/1-332
Chlorella_variabilis[CHLNCRAFT_58828/1-332
Chlorella_sorokiniana[C2E21_8413/1-327
*Auxenochlorella_protothecoides*_(*Chlorella_protothecoides*)[APUTEX25_001631/1-331
*Tetraselmis*_sp._GSL018[TSPGSL018_26319/1-371
*Micromonas_pusilla*_(strain_CCMP1545)[MCPUCRAFT_47615/1-328
*Micromonas_commoda*_(strain_RCC299/NOUM17/CCMP2709)[MCPUN_62224/1-149
*Micromonas_commoda*_(strain_RCC299/NOUM17/CCMP2709)[MCPUN_105899/1-283
Ostreococcus_tauri[BE221DRAFT_194138/1-242
Bathycoccus_prasinos[Bathy10g00200/1-188
Bathycoccus_prasinos[Bathy07g01820/1-312
Gonium_pectoralis[GPECTOR_69g440/1-516
Tettrabaena_socialis[TSOC_008198/1-430
Chlamydomonas_reinhardtii[CHLRE_05g235700v5/1-429
Chlamydomonas_reinhardtii[CHLRE_02g105200v5/1-462
Chlamydomonas_eustigma[CEUSTIGMA_g11715.1/1-345
Klebsomidium_nitens[KFL_001110040/1-305
Chara_braunii[CBR_g3540/1-391
*Phycomitrella_patens*_subsp._patens[PHYPA_022478/1-333
*Phycomitrella_patens*_subsp._patens[PHYPA_018982/1-334
Wollemia_nobilis[WA/1-408
Araucaria_cunninghamii[WA/1-406
Picea_sitchensis[NA][A9NUE1/1-430
Picea_sitchensis[NA][A9P228/1-326
Amborella_trichopoda[AMTR_s00007p00225690/1-214
Amborella_trichopoda[AMTR_s00062p00198130/1-320

133 HVVLIIDG EAYAE EASSAHVASDHGPASECGA RLFVSKFVE 173
71LPTCISFVSQSIN 84
125HGS DTVNSTSMCEFGCQSINQ 146
117 ---GKKTKEAIFGLPAPLKSEIKPLKDTDVCEKAAAKVHD 154
107PAAALQAQVQPRHSSKPNVDTPCLMYVASKLKE 140
101 GASPLARFAAQPPAPRRARSEGRNDMGCPCLMFVVGKVKD 141
101 ---GAPPPAF AAPAKPRRARAAPRNDLTCLCLMFVVSVKVE 138
115GAAPPLSKTARFGVGQQDNFNCPICKMLLITLKQ 148
118PGSSAAQLLGVDAAKGSTPWDCPVCKMVI AFVE 151
119HGGKDLAQVTPFDCPVCRVAQMFVA 144
113VGTAVGVNSANGPFDCPMCRMVALTL LQ 140
124 -FDVELIVLNR ESISSRSHAAIPEELLQDSLLMVEGF FE 163
135 PKDPSRMSRFLSRVAVAPAPRRPGPADLCKKFAVESLHA 175
35ESCRATLT E LRT 48
137 PLEVPA SLREVIERGSAVDAPAPPSTDLCKCEFGVESLHA 177
138 LAIEGVERAGMVDVRRRARGDVAD DSTADVMAT ELLAN 178
92CREIVGT YEE 101
95HEKNENK EEECLSF LHNATE 114
109GDTLAQVPM LTKPALRVHDGAECAMK FVVGRVKA 143
87GAAVLEMTPLNRPAIRVHDNTECAMCKMVVNHVKS 121
146 ASRAIATAASHGHGGSSSTSGGATTMQCATCRHVVESVKA 186
110GAVLAQVPELNKPSLVVRDSTQCLKYVVT LVR 144
107GLKLQSTPHHNQVLQVRVSGGGLCALCSFVMDQM KI 142
115PVPPLFALAKHVL AGKEEQDLOQYAATSLIA 145
175 LSSPRSP LFSPMNHPPPHRRPRHRHSKRRCKVCEAF AAQALR 215
125MPPAIKAFQDEKKKTVQD L ATDAL T 151
127TPPAIKAFQDEKKT VVCEALATDAL T 153
191 -NGNNYGSQIKLWNWTEILPLDVHDDTTCAVCEQFVEEAVY 230
189 -NGNNYDSQIKLWNWTEISPLDVHDDTTCAVCEQFVEEAVY 228
211 NGNNNNDLQYKLNMG RNTPSLDAGDDTTVMCEQFIEEATY 251
211 NGNNNNDLQYKLGMRNTPSLDAGDDTTVMCEQFIEEATY 251
53CEQFASEAF E 62
169 -PGYGETISNNERNIQLNF GSYNKLSSMLRET LLEVANM 207



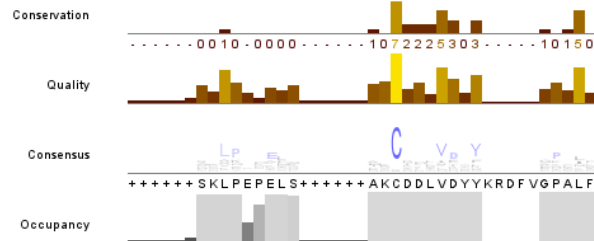
Chloropocon_primus[A3770_07P47130/1-390
Chloropocon_primus[A3770_02p14820/1-272
Chloropocon_primus[A3770_04p29840/1-218
Chloropocon_primus[A3770_04p29830/1-216
Tetrademus_obliquus[BQ4739_LOCUS1502/1-1-209
Raphidocelis_subcapitata[Rauc_10640/1-361
Monoraphidium_neglectum[MNEG_12603/1-188
*Coccomyxa_subellipsoidea*_(strain_C-169)[COGSUBDRAFT_45864/1-332
Chlorella_variabilis[CHLNCRAFT_58828/1-332
Chlorella_sorokiniana[C2E21_8413/1-327
*Auxenochlorella_protothecoides*_(*Chlorella_protothecoides*)[APUTEX25_001631/1-331
*Tetraselmis*_sp._GSL018[TSPGSL018_26319/1-371
*Micromonas_pusilla*_(strain_CCMP1545)[MCPUCRAFT_47615/1-328
*Micromonas_commoda*_(strain_RCC299/NOUM17/CCMP2709)[MCPUN_62224/1-149
*Micromonas_commoda*_(strain_RCC299/NOUM17/CCMP2709)[MCPUN_105899/1-283
Ostreococcus_tauri[BE221DRAFT_194138/1-242
Bathycoccus_prasinos[Bathy10g00200/1-188
Bathycoccus_prasinos[Bathy07g01820/1-312
Gonium_pectoralis[GPECTOR_69g440/1-516
Tettrabaena_socialis[TSOC_008198/1-430
Chlamydomonas_reinhardtii[CHLRE_05g235700v5/1-429
Chlamydomonas_reinhardtii[CHLRE_02g105200v5/1-462
Chlamydomonas_eustigma[CEUSTIGMA_g11715.1/1-345
Klebsomidium_nitens[KFL_001110040/1-305
Chara_braunii[CBR_g3540/1-391
*Phycomitrella_patens*_subsp._patens[PHYPA_022478/1-333
*Phycomitrella_patens*_subsp._patens[PHYPA_018982/1-334
Wollemia_nobilis[WA/1-408
Araucaria_cunninghamii[WA/1-406
Picea_sitchensis[NA][A9NUE1/1-430
Picea_sitchensis[NA][A9P228/1-326
Amborella_trichopoda[AMTR_s00007p00225690/1-214
Amborella_trichopoda[AMTR_s00062p00198130/1-320

174 SWYNTLAKNLAEVK- NNDGKEPPSLTYNEDLEDMIQNL TS 213
85 NLIANGGIVGCSALC- 103
147 FLSKPENDNKT LAFAGVC- 165
155 FLSQPGEIDKIVNFIEEV- 173
141 QVNNPVTQADIRDSALAA- 159
142 GLSDPVTREEIRDKTAAC- 160
139 TLSDPITRQAVHDKTRAAC- 157
148 ELKNPESEKEMIDRAHOAC- 167
152 RLQDREERQQIEADMRAAC- 170
145 RLQDKEARAQVEGVMRAC- 164
141 RFKDPKVRSEMH TGF LQAAC- 159
164 RWTEMMSLQLNRATVQAGGDKPPAI TYRNDMEAMVGEFC- 202
178 AVSSPDAITQLIAQADLD- 195
47 MVDKTAKKQGRENAVTDAMEAICED 71
178 AITSNATVQSLNNEAEKACE- 197
179 EIKSNSTIN FVEGEVDALC- 197
102KIERLMYKYAKTVE- 115
115 KVRSKEFRARLYKVVD EMO- 134
144 AINDSGTLEKIKEVALQLC- 162
122 AINDSQTMEKIKEVALQAAC- 140
187 AAEARGEDGEGGLNSPAGAHMAARAAAC- 213
145 AVNSTATLEKIEQAALQAAC- 163
143 ALNSTSVQQT IMEKAQEIC- 161
146 YLASNQTQQQIMTVAHTAC- 164
216 YLGQNETVAEIVSLAHKAC- 234
152 YLENNKTREEIVIALHLGC- 170
154 YLDNNRTREEIVVALHLAC- 172
231 YVDQNKTRSEILSALHQTC- 249
229 YVDQNKTRSEILSALHQTC- 247
252 YASQNTTQSEVL AALHQT C- 270
252 YASQNATQSEVL AALHQT C- 270
63 YLGNNQTQTDI IKTLHQVC- 81
208 AYQRQEQSEVLNEK- 221



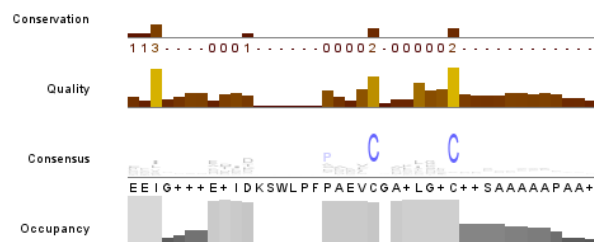
Chloropocon_primus[A3770_07P47130/1-390
Chloropocon_primus[A3770_02p14820/1-272
Chloropocon_primus[A3770_04p29840/1-218
Chloropicon_primus[A3770_04p29830/1-216
Tetrademus_obliquus[BQ4739_LOCUS1502/1-209
Raphidocelia_subcapitata[Rauc_10640/1-361
Monoraphidium_neglectum[MNEG_12603/1-188
Coccomyxa_subellipsoidea (strain_C-169)[COC SUBDRAFT_45864/1-332
Chlorella_variabilis[CHLNC DRAFT_58828/1-332
Chlorella_sorokiniana[C2E21_8413/1-327
Auxenochlorella_protothecoides (Chlorella_protothecoides)[APUTEX25_001631/1-331
Tetraselmis_sp. GSL018[TSPGSL018_26319/1-371
Micromonas_pusilla (strain_CCMP1545)[MCPUC DRAFT_47615/1-328
Micromonas_commoda (strain_RCC299/NOUM17/CCMP2709)[MCPUN_62224/1-149
Micromonas_commoda (strain_RCC299/NOUM17/CCMP2709)[MCPUN_105899/1-283
Ostreococcus_taui[BE221DRAFT_194138/1-242
Bathycoccus_prasinos[Bathy10g00200/1-188
Bathycoccus_prasinos[Bathy07g01820/1-312
Gonium_pectorale[GPECTOR_69g440/1-516
Tetrabaena_socialis[TSOC_008198/1-430
Chlamydomonas_reinhardtii[CHLRE_05g235700v5/1-429
Chlamydomonas_reinhardtii[CHLRE_02g105200v5/1-462
Chlamydomonas_eustigma[CEUSTIGMA_g11715.t1/1-345
Kleboormidium_nitens[KFL_001110040/1-305
Chara_braunii[CBR_g3540/1-391
Physcomitrella_patens_subsp._patens[PHYPA_022478/1-333
Physcomitrella_patens_subsp._patens[PHYPA_018982/1-334
Wollemia_nobilis[WA/1-408
Araucaria_cunninghamii[WA/1-406
Picea_sitchensis[NA]A9NUE1/1-430
Picea_sitchensis[NA]A9P228/1-326
Amborella_trichopoda[AMTR_s00007p00225690/1-214
Amborella_trichopoda[AMTR_s00062p00198130/1-320

214 DVFLKDTKVDTKAIL.....PFDEKTMVH.....KRTIA 243
104.....GMLPNQTLG.....EVCALLCIV.....GIDMF 127
176.....PLFS-PELQ.....PKCAQDIMMF.....GPEVI 188
107.....ELVPSDEH.....DKCVSTITGF.....GPMVL 196
160.....AALPEGMMR.....DACTDFVEQY.....GEQWK 183
161.....ATLPEGPMR.....DTCDAWGRQY.....EDSIF 184
158.....GVMPEGAVR.....DACIQWADQY.....GGTLG 181
168.....KTLP-IDWQ.....APCTAYVDQF.....GEQLF 190
171.....DNLP-PEAK.....ARCLDDVTNL.....FVALD 193
165.....ASNLP-PEGK.....AKCSDVTDL.....FSALD 188
160.....NDLE-PAKR.....PKCVTDVDAL.....FQAVE 182
203.....NSAPI NSDRVAGFVQPAKCEIMKVWKRDFVGHFLS 237
196.....KYAAQFELD.....AECEAAVEKY.....GPSAL 219
72 MYNFRTYAYPPQMQ.....KGCRTIMDRH.....EEEIE 101
198.....KYAAAFDMS.....ATCEAAIETY.....GPELV 221
198.....AALG-PELA.....HQDAVLEPY.....VPALL 220
116.....KNKKKKGEDS.....EDSGGENEAQ.....EEAMM 140
135.....RVFPDDEGEEDDKKIAECEALGKTY.....AKRAI 164
163.....GGLP-NELA.....TSCDFVNSY.....VPLIA 185
141.....AGLP-QELS.....DSCTDFVNAY.....EPLMA 163
214.....AGLP-GPLA.....AACSDVDRR.....SAILL 236
164.....SALP-AELA.....STCTDFVNTY.....APMIA 186
162.....HSLP-ADIG.....TSCIDFVQTY.....EPLVT 184
165.....LHVKKPELR.....AQCDTAVDEY.....VPQLL 188
235.....QLLKGETS.....EECNLSLVEVY.....AQDPL 258
171.....SGLR-ELS.....KSCDLLVDLY.....SARMN 192
173.....AQMK-ELS.....KQCDLLVDVY.....TPRMM 104
250.....SKLK-MFS.....TECDLSLVDY.....ASLFF 271
248.....SKLK-MFS.....TECDLSLVDY.....ASLFF 269
271.....SKLG-AFG.....TKCKTMVDY.....APIIF 292
271.....SKLG-VFG.....TKCKTMVDY.....APIIF 292
82.....SSMY-SFK.....HQCTSLVDY.....LPMIF 103
222.....DHSSTVLLRF.....NEKAS 236

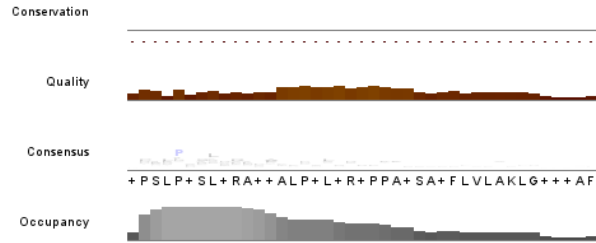


Chloropocon_primus[A3770_07P47130/1-390
Chloropocon_primus[A3770_02p14820/1-272
Chloropocon_primus[A3770_04p29840/1-218
Chloropicon_primus[A3770_04p29830/1-216
Tetrademus_obliquus[BQ4739_LOCUS1502/1-209
Raphidocelia_subcapitata[Rauc_10640/1-361
Monoraphidium_neglectum[MNEG_12603/1-188
Coccomyxa_subellipsoidea (strain_C-169)[COC SUBDRAFT_45864/1-332
Chlorella_variabilis[CHLNC DRAFT_58828/1-332
Chlorella_sorokiniana[C2E21_8413/1-327
Auxenochlorella_protothecoides (Chlorella_protothecoides)[APUTEX25_001631/1-331
Tetraselmis_sp. GSL018[TSPGSL018_26319/1-371
Micromonas_pusilla (strain_CCMP1545)[MCPUC DRAFT_47615/1-328
Micromonas_commoda (strain_RCC299/NOUM17/CCMP2709)[MCPUN_62224/1-149
Micromonas_commoda (strain_RCC299/NOUM17/CCMP2709)[MCPUN_105899/1-283
Ostreococcus_taui[BE221DRAFT_194138/1-242
Bathycoccus_prasinos[Bathy10g00200/1-188
Bathycoccus_prasinos[Bathy07g01820/1-312
Gonium_pectorale[GPECTOR_69g440/1-516
Tetrabaena_socialis[TSOC_008198/1-430
Chlamydomonas_reinhardtii[CHLRE_05g235700v5/1-429
Chlamydomonas_reinhardtii[CHLRE_02g105200v5/1-462
Chlamydomonas_eustigma[CEUSTIGMA_g11715.t1/1-345
Kleboormidium_nitens[KFL_001110040/1-305
Chara_braunii[CBR_g3540/1-391
Physcomitrella_patens_subsp._patens[PHYPA_022478/1-333
Physcomitrella_patens_subsp._patens[PHYPA_018982/1-334
Wollemia_nobilis[WA/1-408
Araucaria_cunninghamii[WA/1-406
Picea_sitchensis[NA]A9NUE1/1-430
Picea_sitchensis[NA]A9P228/1-326
Amborella_trichopoda[AMTR_s00007p00225690/1-214
Amborella_trichopoda[AMTR_s00062p00198130/1-320

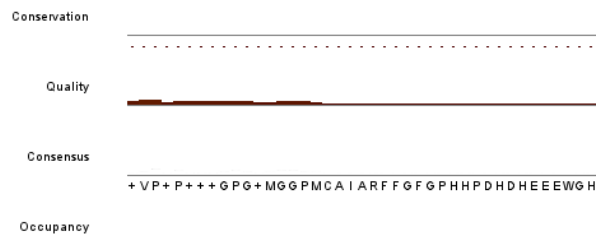
244 EQAMRHLNDMTKSWLPFSNQVGLMTQSPLRG..... 276
128 VKI-LSKVLDL.....PFYAC-ELLTVCHPTDGGAAANITD 160
189 DFI-MQVTSD.....PRDTC-NKLKF..... 208
197 NYL-VGLTAN.....PQQAQ-AEISLC..... 216
184 GSSPAAGQQQQ.....QQQQ-QLLMMC..... 204
185 EYI-DTTE.....PADL-AALGSVLTRLLA.... 210
182 LTR-RGTE..... 188
191 AYI-ETLD.....ATTIC-TNFGALPSQLYSNMTV- 219
194 HLL-RDVD.....PEGVC-RVAEFTPSGAASA.... 219
198 GLL-HDVD.....AEGVC-AVDFCAQPGGS.... 212
183 NIM-DDLD.....PDMAC-EVAQFCPPRALGAGTGAA 212
238 NTVGHGQLPEK.....RAAIRDVMVG..... 260
220 TKA-SAVLAD.....PARVC-AELEMCPAENPAASAARA 251
102 TALWRGDENL.....VEFICGRPKGA..... 123
222 NAAKGYIED.....AHKVC-SELGMC..... 241
221 NVL-AEK.....LRDAQ-AKIGAC..... 237
141 DDI.....FMKVC..... 148
165 EYV-KTHEEET.....DEFVC-EKLNLCADDATVFP... 193
188 ELV-EDMD.....PDTVC-ALVGVMESLVAHP... 212
164 EFV-EDMD.....SDTVC-GLLGVCAESAAAIAP... 190
237 LAVGGTAARD.....TAEAC-GLLGMCGLSVTTAGAAAP 270
187 QLI-ESMD.....ADTVC-GLAGVMEAAAAVFP... 213
185 QFI-ESVD.....SAQMC-LLIGACLDGAVLKAP... 211
189 AVM-QTIT.....PAELC-HILSF..... 206
259 SVM-ASIT.....PEQFC-HLLKFCDGDDNPAAAAAD 288
193 EQL-ENIT.....PQEFQ-QMTKMC..... 210
195 EQL-GNIT.....PQSFC-EKTRMC..... 212
272 MEL-ETVK.....PKEFC-QKISY..... 289
270 MEV-ETVK.....PKEFC-QKISY..... 287
293 LEI-ATIS.....PKEFC-QKISY..... 310
293 LEI-ATIS.....LATIC-LYSKIV..... 310
104 SEI-AMIN.....PEGLC-AKVNLC..... 121
237 GNI..... 239



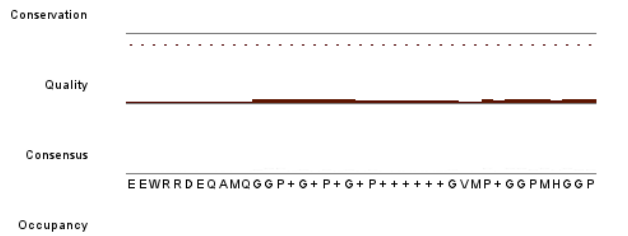
Chloropocon_primus[A3770_07P47130/1-390
Chloropocon_primus[A3770_02p14820/1-272
Chloropocon_primus[A3770_04p29840/1-218
Chloropicon_primus[A3770_04p29830/1-216
Tetrademus_obliquus[BQ4739_LOCUS15021/1-209
Raphidocelia_subcapitata[Rscu_10640/1-361
Monoraphidium_neglectum[MNEG_12603/1-188
*Coccomyxa_subellipsoidea*_(strain_C-169)[COCSUBDRAFT_45864/1-332
Chlorella_variabilis[CHLNCRAFT_58828/1-332
Chlorella_sorokiniana[C2E21_8413/1-327
*Auxenochlorella_protothecoides*_(*Chlorella_protothecoides*)[APUTEX25_001631/1-331
*Tetraselmis*_sp.GSL018[TSPGSL018_26319/1-371
*Micromonas_puilla*_(strain_CCMP1545)[MCPUCRAFT_47615/1-328
*Micromonas_commoda*_(strain_RCC299/NOUM17/CCMP2709)[MCPUN_62224/1-149
*Micromonas_commoda*_(strain_RCC299/NOUM17/CCMP2709)[MCPUN_105899/1-283
Ostreococcus_taui[BE221DRAFT_194138/1-242
Bathycoccus_prasinus[Bathy10g00200/1-188
Bathycoccus_prasinus[Bathy07g01820/1-312
Gonium_pectorale[GPECTOR_69g440/1-516
Tetrabaena_socialis[TSOC_008198/1-430
Chlamydomonas_reinhardtii[CHLRE_05g235700v5/1-429
Chlamydomonas_reinhardtii[CHLRE_02g105200v5/1-462
Chlamydomonas_eustigma[CEUSTIGMA_g11715.1/1-345
Klebsormidium_nitens[KFL_001110040/1-305
Chara_braunii[CBR_g3540/1-391
Phycomitrella_patens_subsp._patens[PHYPA_022478/1-333
Phycomitrella_patens_subsp._patens[PHYPA_018982/1-334
Wollemia_nobilis[NA/1-408
Araucaria_cunninghamii[NA/1-406
Picea_sitchensis[NA|A9NUE1/1-430
Picea_sitchensis[NA|A9P228/1-326
Amborella_trichopoda[AMTR_s00007p00225690/1-214
Amborella_trichopoda[AMTR_s00062p00198130/1-320



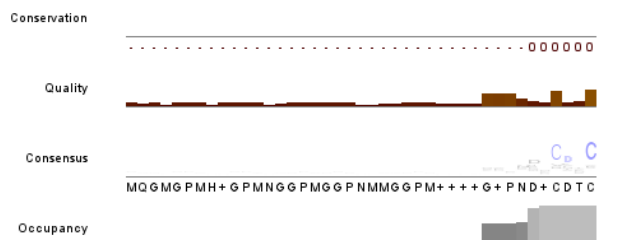
Chloropocon_primus[A3770_07P47130/1-390
Chloropocon_primus[A3770_02p14820/1-272
Chloropocon_primus[A3770_04p29840/1-218
Chloropicon_primus[A3770_04p29830/1-216
Tetrademus_obliquus[BQ4739_LOCUS15021/1-209
Raphidocelia_subcapitata[Rscu_10640/1-361
Monoraphidium_neglectum[MNEG_12603/1-188
*Coccomyxa_subellipsoidea*_(strain_C-169)[COCSUBDRAFT_45864/1-332
Chlorella_variabilis[CHLNCRAFT_58828/1-332
Chlorella_sorokiniana[C2E21_8413/1-327
*Auxenochlorella_protothecoides*_(*Chlorella_protothecoides*)[APUTEX25_001631/1-331
*Tetraselmis*_sp.GSL018[TSPGSL018_26319/1-371
*Micromonas_puilla*_(strain_CCMP1545)[MCPUCRAFT_47615/1-328
*Micromonas_commoda*_(strain_RCC299/NOUM17/CCMP2709)[MCPUN_62224/1-149
*Micromonas_commoda*_(strain_RCC299/NOUM17/CCMP2709)[MCPUN_105899/1-283
Ostreococcus_taui[BE221DRAFT_194138/1-242
Bathycoccus_prasinus[Bathy10g00200/1-188
Bathycoccus_prasinus[Bathy07g01820/1-312
Gonium_pectorale[GPECTOR_69g440/1-516
Tetrabaena_socialis[TSOC_008198/1-430
Chlamydomonas_reinhardtii[CHLRE_05g235700v5/1-429
Chlamydomonas_reinhardtii[CHLRE_02g105200v5/1-462
Chlamydomonas_eustigma[CEUSTIGMA_g11715.1/1-345
Klebsormidium_nitens[KFL_001110040/1-305
Chara_braunii[CBR_g3540/1-391
Phycomitrella_patens_subsp._patens[PHYPA_022478/1-333
Phycomitrella_patens_subsp._patens[PHYPA_018982/1-334
Wollemia_nobilis[NA/1-408
Araucaria_cunninghamii[NA/1-406
Picea_sitchensis[NA|A9NUE1/1-430
Picea_sitchensis[NA|A9P228/1-326
Amborella_trichopoda[AMTR_s00007p00225690/1-214
Amborella_trichopoda[AMTR_s00062p00198130/1-320



Chloropocon_primus[A3770_07P47130/1-390
Chloropocon_primus[A3770_02p14820/1-272
Chloropocon_primus[A3770_04p29840/1-218
Chloropicon_primus[A3770_04p29830/1-216
Tetrademus_obliquus[BQ4739_LOCUS15021/1-209
Raphidocelia_subcapitata[Rsuc_10640/1-361
Monoraphidium_neglectum[MNEG_12603/1-188
Coccomyxa_subellipsoidea_(strain_C-169)[COC SUBDRAFT_45864/1-332
Chlorella_variabilis[CHLNC DRAFT_58828/1-332
Chlorella_sorokiniana[C2E21_8413/1-327
Auxenochlorella_prothothecoides_(Chlorella_prothothecoides)[APUTEX25_001631/1-331
Tetraselmis_sp.GSL018[TSPGSL018_26319/1-371
Micromonas_pusilla(strain_CCMP1545)[MCPUC DRAFT_47615/1-328
Micromonas_commoda(strain_RCC299/NOLM17/CCMP2709)[MCPUN_62224/1-149
Micromonas_commoda_(strain_RCC299/NOLM17/CCMP2709)[MCPUN_105899/1-283
Ostreococcus_taui[BE221DRAFT_194138/1-242
Bathycoccus_prasinus[Bathy10g00200/1-188
Bathycoccus_prasinus[Bathy07g01820/1-312
Gonium_pectorale[GPECTOR_69g440/1-516
Tetrabaena_socialis[TSOC_008198/1-430
Chlamydomonas_reinhardtii[CHLRE_05g235700v5/1-429
Chlamydomonas_reinhardtii[CHLRE_02g105200v5/1-462
Chlamydomonas_eustigma[CEUSTIGMA_g11715.1/1-345
Klebsormidium_nitens[KFL_001110040/1-305
Chara_braunii[CBR_g3540/1-391
Physcomitrella_patens_subsp._patens[PHYPA_022478/1-333
Physcomitrella_patens_subsp._patens[PHYPA_018982/1-334
Wollemia_nobilis[NA/1-408
Araucaria_cunninghamii[NA/1-406
Picea_sitchensis[NA|A9NUE1/1-430
Picea_sitchensis[NA|A9P228/1-326
Amborella_trichopoda[AMTR_s00007p00225690/1-214
Amborella_trichopoda[AMTR_s00062p00198130/1-320

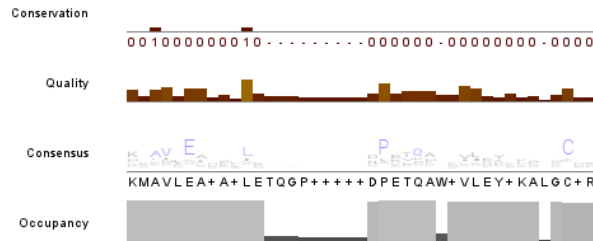


Chloropocon_primus[A3770_07P47130/1-390
Chloropocon_primus[A3770_02p14820/1-272
Chloropocon_primus[A3770_04p29840/1-218
Chloropicon_primus[A3770_04p29830/1-216
Tetrademus_obliquus[BQ4739_LOCUS15021/1-209
Raphidocelia_subcapitata[Rsuc_10640/1-361
Monoraphidium_neglectum[MNEG_12603/1-188
Coccomyxa_subellipsoidea_(strain_C-169)[COC SUBDRAFT_45864/1-332
Chlorella_variabilis[CHLNC DRAFT_58828/1-332
Chlorella_sorokiniana[C2E21_8413/1-327
Auxenochlorella_prothothecoides_(Chlorella_prothothecoides)[APUTEX25_001631/1-331
Tetraselmis_sp.GSL018[TSPGSL018_26319/1-371
Micromonas_pusilla(strain_CCMP1545)[MCPUC DRAFT_47615/1-328
Micromonas_commoda(strain_RCC299/NOLM17/CCMP2709)[MCPUN_62224/1-149
Micromonas_commoda_(strain_RCC299/NOLM17/CCMP2709)[MCPUN_105899/1-283
Ostreococcus_taui[BE221DRAFT_194138/1-242
Bathycoccus_prasinus[Bathy10g00200/1-188
Bathycoccus_prasinus[Bathy07g01820/1-312
Gonium_pectorale[GPECTOR_69g440/1-516
Tetrabaena_socialis[TSOC_008198/1-430
Chlamydomonas_reinhardtii[CHLRE_05g235700v5/1-429
Chlamydomonas_reinhardtii[CHLRE_02g105200v5/1-462
Chlamydomonas_eustigma[CEUSTIGMA_g11715.1/1-345
Klebsormidium_nitens[KFL_001110040/1-305
Chara_braunii[CBR_g3540/1-391
Physcomitrella_patens_subsp._patens[PHYPA_022478/1-333
Physcomitrella_patens_subsp._patens[PHYPA_018982/1-334
Wollemia_nobilis[NA/1-408
Araucaria_cunninghamii[NA/1-406
Picea_sitchensis[NA|A9NUE1/1-430
Picea_sitchensis[NA|A9P228/1-326
Amborella_trichopoda[AMTR_s00007p00225690/1-214
Amborella_trichopoda[AMTR_s00062p00198130/1-320



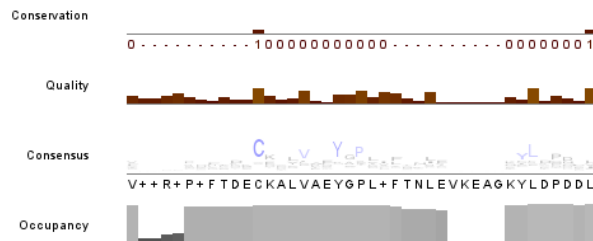
Chloropocon_primus[A3770_07P47130/1-390
Chloropocon_primus[A3770_02p14820/1-272
Chloropocon_primus[A3770_04p29840/1-218
Chloropocon_primus[A3770_04p29830/1-216
Tetrademus_obliquus[BQ4739_LOCUS15021/1-209
Raphidocelis_subcapitata[Rscu_10640/1-361
Monoraphidium_neglectum[MNEG_12603/1-188
Coccomyxa_subellipsoidea (strain_C-169)[COC SUBDRAFT_45864/1-332
Chlorella_variabilis[CHLNC DRAFT_58828/1-332
Chlorella_sorokiniana[C2E21_8413/1-327
Auxenochlorella_protothecoides (Chlorella_protothecoides)[APUTEX25_001631/1-331
*Tetraselmis*_sp. GSL018[TSPGSL018_26319/1-371
Micromonas_pusilla (strain_CCMP1545)[MCPUC DRAFT_47615/1-328
Micromonas_commoda (strain_RCC299/NOLM17/CCMP2709)[MCPUN_62224/1-149
Micromonas_commoda (strain_RCC299/NOLM17/CCMP2709)[MCPUN_105899/1-283
Ostreococcus_tauri[BE221 DRAFT_194138/1-242
Bathycoccus_prasinos[Bathy10g00200/1-188
Bathycoccus_prasinos[Bathy07g01820/1-312
Gonium_pectorale[GPECTOR_69g440/1-516
Tetrabaena_socialis[TSOC_008198/1-430
Chlamydomonas_reinhardtii[CHLRE_05g235700v5/1-429
Chlamydomonas_reinhardtii[CHLRE_02g105200v5/1-462
Chlamydomonas_eustigma[CEUSTGMA_g117151/1-345
Klebsormidium_nitens[KFL_001110040/1-305
Chara_braunii[CBR_g3540/1-391
Physcomitrella_patens_subsp._patens[PHYPA_022478/1-333
Physcomitrella_patens_subsp._patens[PHYPA_018982/1-334
Wollemia_nobilis[NA/1-408
Araucaria_cunninghamii[NA/1-406
Picea_sitchensis[NA]A9NUE1/1-430
Picea_sitchensis[NA]A9P228/1-326
Amborella_trichopoda[AMTR_s00007p00225690/1-214
Amborella_trichopoda[AMTR_s00062p00198130/1-320

287 KALIEEIEFDLATQGPSLKVLP I EKRAWG ILTDMCG . KTKY 326
185 PCGDGELIFA IQ D EGNVLGSGQVFGSLQPNV 216
.
205 G C A S 208
249 KAVVT E M H S A L A N P E L E A . Q V E Q Y A K A . V C D S 278
.
251 K M V V T E A A A I L G N L D T Q K . Q I L E Y A K E . A C Q A 280
255 K T I V M E A A A I L Q D P K T Q A . E L L E Y A K E . G C N I 284
249 K T I V M E A A A I L Q D P K T Q A . E L L E Y A K Q . G C T I 278
248 K V V V L Q A V A I I Q V G A E A V N G A N P K T Q A . D I I E Y A K E . S C S M 286
276 R A T V D A L A F E L R T Q G P L Q E A R G R K R Q V W D T L D F I C M . K L H Y 315
280 K G A L G E A R A K L M K K G G P E R M T R S V L K S S V A 309
130 E E A V N S K P M E I V . 141
252 D S K K T R R R E R M K . 282
238 . 242
154 K G I D L E T Y M Q I E . 180
240 E Y G A S K L A V A M N . 269
371 K M A V I E A H S L I S . 400
302 K M A V I E A H S L V S . 331
325 K M A V A E V S L L S . 354
313 K M A V I E A H S L V S . 342
244 K M A V I E A H S L V T . 273
224 H F L V L E L R F K L E . 253
318 R Y A V V K A R K R L E . 347
226 Q F A I L E I K I K L E . 255
227 Q F F I L E I K I K L Q . 256
307 R A A M S E V K A E L E . 336
305 R A A M S E V K A E L E . 334
329 E S A M L E I E T H L K . 358
.
141 H R A V L E I L M K L K . 170
245 H R A I D E I E K D L R . 274



Chloropocon_primus[A3770_07P47130/1-390
Chloropocon_primus[A3770_02p14820/1-272
Chloropocon_primus[A3770_04p29840/1-218
Chloropocon_primus[A3770_04p29830/1-216
Tetrademus_obliquus[BQ4739_LOCUS15021/1-209
Raphidocelis_subcapitata[Rscu_10640/1-361
Monoraphidium_neglectum[MNEG_12603/1-188
Coccomyxa_subellipsoidea (strain_C-169)[COC SUBDRAFT_45864/1-332
Chlorella_variabilis[CHLNC DRAFT_58828/1-332
Chlorella_sorokiniana[C2E21_8413/1-327
Auxenochlorella_protothecoides (Chlorella_protothecoides)[APUTEX25_001631/1-331
*Tetraselmis*_sp. GSL018[TSPGSL018_26319/1-371
Micromonas_pusilla (strain_CCMP1545)[MCPUC DRAFT_47615/1-328
Micromonas_commoda (strain_RCC299/NOLM17/CCMP2709)[MCPUN_62224/1-149
Micromonas_commoda (strain_RCC299/NOLM17/CCMP2709)[MCPUN_105899/1-283
Ostreococcus_tauri[BE221 DRAFT_194138/1-242
Bathycoccus_prasinos[Bathy10g00200/1-188
Bathycoccus_prasinos[Bathy07g01820/1-312
Gonium_pectorale[GPECTOR_69g440/1-516
Tetrabaena_socialis[TSOC_008198/1-430
Chlamydomonas_reinhardtii[CHLRE_05g235700v5/1-429
Chlamydomonas_reinhardtii[CHLRE_02g105200v5/1-462
Chlamydomonas_eustigma[CEUSTGMA_g117151/1-345
Klebsormidium_nitens[KFL_001110040/1-305
Chara_braunii[CBR_g3540/1-391
Physcomitrella_patens_subsp._patens[PHYPA_022478/1-333
Physcomitrella_patens_subsp._patens[PHYPA_018982/1-334
Wollemia_nobilis[NA/1-408
Araucaria_cunninghamii[NA/1-406
Picea_sitchensis[NA]A9NUE1/1-430
Picea_sitchensis[NA]A9P228/1-326
Amborella_trichopoda[AMTR_s00007p00225690/1-214
Amborella_trichopoda[AMTR_s00062p00198130/1-320

327 SHSNPHKASEV E D L L E D H G K A L V S L I A V K E A E G G F L D N G 367
217 Y S I D F Q V K T E P S Q Q Q G G E Q W P A G D Y M A E M A V 247
209 L A P N E T 214
.
209 V 209
279 M G A L A D S C K E R I D Q Y A P M A F G M I L A Y L Q P D Q V 310
.
281 F G P N F K D Q C L N Y V E L Y G P L V V N M I V Q Y L K P Q . L 312
285 F A D F K D Q C E Q Y V T L Y G P L V F N M L I S Y L Q P D S V 316
279 F Q D F K D Q C E A Y V T L Y G P L V F N M L I S Y L Q P D F L 310
287 F P D W Q D S C E A Y V T L Y G P I A I N M L V T Y L Q P Q T V 318
318 L R A R A T K A Q E L Q D D L V D E H G S E I V R M V E E G E P W A A V 351
310 R R N K A K A E R E K T S L G A K R R 328
142 . 149
283 I . 283
.
181 . 188
270 A E D R F S D K E Q K E A A A A E V E S K I Y D A V Q T W L D D G E A 305
401 F P N F A Q P C K T Y V A V Y A P L V F S L L E Q Y L V P D Q L 432
332 F S A F A D P C K A Y V D M Y A P L V F T L L E Q Y L V P D T L 363
355 A L P S A Y T P A C R I A V D A Y A P M I F A L V P Q Y V Q P D P V 388
343 F P S F T T A C K S Y V A M Y A P L V F T L L E Q Y L V P D T L 374
274 V G G S A L S E T Q R E Y V D T Y A I E V F K L M D K L L T P D Q L 307
254 M P N L A A Q C T E S I A E Y A P V I F Q S L D A V M D P R L I 285
348 V P K H T D Q C K A F L E S Y S E N F F A N L D V I L D P H G F 379
256 V V N H V D E C K L L V A Q Y G P F V L A N I D K I L D S Q A L 287
257 I T N H V D E C K S L V A Q Y G P I A L A D I D K V L D A Q A V 288
337 V P T Y T K E C K R L V F E Y G P V I L T N M E K Y L D S N D I 368
335 V P T Y T K E C K R L V F E Y G P I L T N M E K Y L D S N D I 366
359 V Q G H E Q C K K L V F E Y G P I L T N L E K Y L D S N D I 360
. 326
171 A Q G F V E K C E D L V F E Y A P L L L I N A E Q F L E T K D I 202
275 V Q D H V K E C K K L V L E Y V P L I L V N L E K Y L K N N D I 306



<i>Chloropocoon_primus</i> A3770_07P47130/1-330	368	MTELFCKSA	MCNDQKSYDKDEL	390
<i>Chloropocoon_primus</i> A3770_02p14820/1-272	248	CQ	GEIG	253
<i>Chloropocoon_primus</i> A3770_04p29840/1-218	215	VALY		218
<i>Chloropocoon_primus</i> A3770_04p29830/1-216				
<i>Tetrademus_obliquus</i> BQ4739_LOCUS15021/1-209				
<i>Raphidocelis_subcapitata</i> [Rauc_10640/1-361	311	CRQM	HFCP	322
<i>Monoraphidium_neglectum</i> [MNEG_12603/1-188				
<i>Coccomyxa_subellipsoidea</i> (strain_C-169) [COC SUBDRAFT_45864/1-332	313	CIDA	GYCPQPTSLNGLMRSV	332
<i>Chlorella_variabilis</i> [CHLNC DRAFT_58828/1-332	317	CTRM	GYCS	324
<i>Chlorella_sorokiniana</i> [C2E21_84131-327	311	CARL	GYCP	318
<i>Auxenochlorella_protothecoides</i> (<i>Chlorella_protothecoides</i>) [APUTEX26_001631/1-331	319	CSEL	GYCPPLSLA	331
<i>Tetraselmis</i> sp. GSL018 [TSPGSL018_26319/1-371	352	SSRLCGTLTGHS		364
<i>Micromonas_pusilla</i> (strain_CCMP1545) [MCPUC DRAFT_47615/1-328				
<i>Micromonas_commoda</i> (strain_RCC299/NOUM17/CCMP2709) [MCPUN_62224/1-149				
<i>Micromonas_commoda</i> (strain_RCC299/NOUM17/CCMP2709) [MCPUN_105899/1-283				
<i>Ostreococcus_tauri</i> [BE221 DRAFT_194138/1-242				
<i>Bathycoccus_prasinos</i> [Bathy10g00200/1-188				
<i>Bathycoccus_prasinos</i> [Bathy07g01820/1-312	306	CEDL	FSC	312
<i>Gonium_pectorale</i> [GPECTOR_63g440/1-516	433	CAQT	GMCPPPPHPRGGEEVPPHP	456
<i>Tetrabasena_socialis</i> [TSOC_008198/1-430	364	CSQT	GMCPPPP	382
<i>Chlamydomonas_reinhardtii</i> [CHLRE_05g235700v5/1-429	389	CVRL	GMCP	396
<i>Chlamydomonas_reinhardtii</i> [CHLRE_03g105200v5/1-462	375	CAQT	GMCPPPP	408
<i>Chlamydomonas_eustigma</i> [CEUS TIGMA_g11715.1/1-345	380	CTQL	GICKP	316
<i>Klebsomidium_nitens</i> [KFL_001110040/1-305	286	CKKL	HVCP	293
<i>Chara_braunii</i> [CBR_g3540/1-391	380	CRRM	GACFSSSM	391
<i>Physcomitrella_patens</i> subsp. <i>patens</i> [PHYPA_022478/1-333	288	CKT	GFCQ	295
<i>Physcomitrella_patens</i> subsp. <i>patens</i> [PHYPA_018982/1-334	289	CKA	GICK	296
<i>Wollemia_nobilis</i> [WA/1-408	369	CSEL	HVCE	376
<i>Araucaria_cunninghamii</i> [WA/1-406	367	CSEL	HVCE	374
<i>Picea_sitchensis</i> [WA] A9P140/1-430	391	CSQI	HVCE	398
<i>Picea_sitchensis</i> [WA] A9P228/1-326				
<i>Amborella_trichopoda</i> [AMTR_s00007p00225690/1-214	203	CAV	HVCK	210
<i>Amborella_trichopoda</i> [AMTR_s00062p00198130/1-320	307	CAML	HVCKDHMILL	320

<i>Chloropocoon_primus</i> [A3770_07P47130/1-3	254	SH	255
<i>Chloropocoon_primus</i> [A3770_02p14820/1-272			
<i>Chloropocoon_primus</i> [A3770_04p29840/1-218			
<i>Chloropocoon_primus</i> [A3770_04p29830/1-216			
<i>Tetrademus_obliquus</i> [BQ4739_LOCUS1502/1-209			
<i>Raphidocelis_subcapitata</i> [Rauc_10640/1-361	323	AEQLLMGVARLVSP	337
<i>Monoraphidium_neglectum</i> [MNEG_12603/1-188			
<i>Coccomyxa_subellipsoidea</i> _(strain_C-169)[COCSUBDRAFT_45864/1-332			
<i>Chlorella_variabilis</i> [CHLNCRAFT_58828/1-332			
<i>Chlorella_sorokiniana</i> [C2E21_8413/1-327			
<i>Auxenochlorella_protothecoides</i> _(<i>Chlorella_protothecoides</i>)[APUTEX25_001631/1-331			
<i>Tetraselmis</i> _sp.[GSL018][TSPGSL018_26319/1-371			
<i>Micromonas_pusilla</i> _(strain_CCMP1545)[MCPUCRAFT_47615/1-328			
<i>Micromonas_commoda</i> _(strain_RCC299/NOUM17/CCMP2709)[MCPUN_62224/1-149			
<i>Micromonas_commoda</i> _(strain_RCC299/NOUM17/CCMP2709)[MCPUN_105899/1-283			
<i>Ostreococcus_tauri</i> [BE221DRAFT_194138/1-242			
<i>Bathycoccus_prasinos</i> [Bathy10g00200/1-188			
<i>Bathycoccus_prasinos</i> [Bathy07g01820/1-312			
<i>Gonium_pectorale</i> [GPECTOR_63g440/1-516	457	LSHPHPHGPLG-WLLDGLVGLAQRIG	482
<i>Tetrahena_socialis</i> [TSOC_008198/1-430	383	MEEP-PTLGC-W-FGALLHRVGGFFAHL	408
<i>Chlamydomonas_reinhardtii</i> [CHLRE_05g235700v5/1-429	397	VLSAVEQLVLCGKG	410
<i>Chlamydomonas_reinhardtii</i> [CHLRE_02g105200v5/1-462	409	LQKASAAAAALQPPPHPSWGCW-FHDVMRNVYNNFAR	444
<i>Chlamydomonas_eustigma</i> [CEUSTIGMA_g11715.t1/1-345	317	SFIYLIKQAQEA	330
<i>Klebsomidium_nitens</i> [KFL_001110040/1-305	294	RAVSLQEAIRGV	305
<i>Chara_braunii</i> [CBR_g3540/1-391			
<i>Phycomitrella_patens_subsp._patens</i> [PHYPA_022478/1-333	298	SRICPQKAKGWTET	309
<i>Phycomitrella_patens_subsp._patens</i> [PHYPA_018982/1-334	297	SRVCPQKAKAWTET	310
<i>Wollemia_nobilis</i> [WA/1-408	377	SAAGKLENKINDRV	360
<i>Araucaria_cunninghamii</i> [WA/1-406	375	SHAGKLENNMDRV	388
<i>Picea_sitchensis</i> [WA/A9NUE1/1-430	399	NATRD	403
<i>Picea_sitchensis</i> [WA/A9P228/1-326			
<i>Amborella_trichopoda</i> [AMTR_s00007p00225690/1-214	211	AF CY	214
<i>Amborella_trichopoda</i> [AMTR_s00062p00198130/1-320			

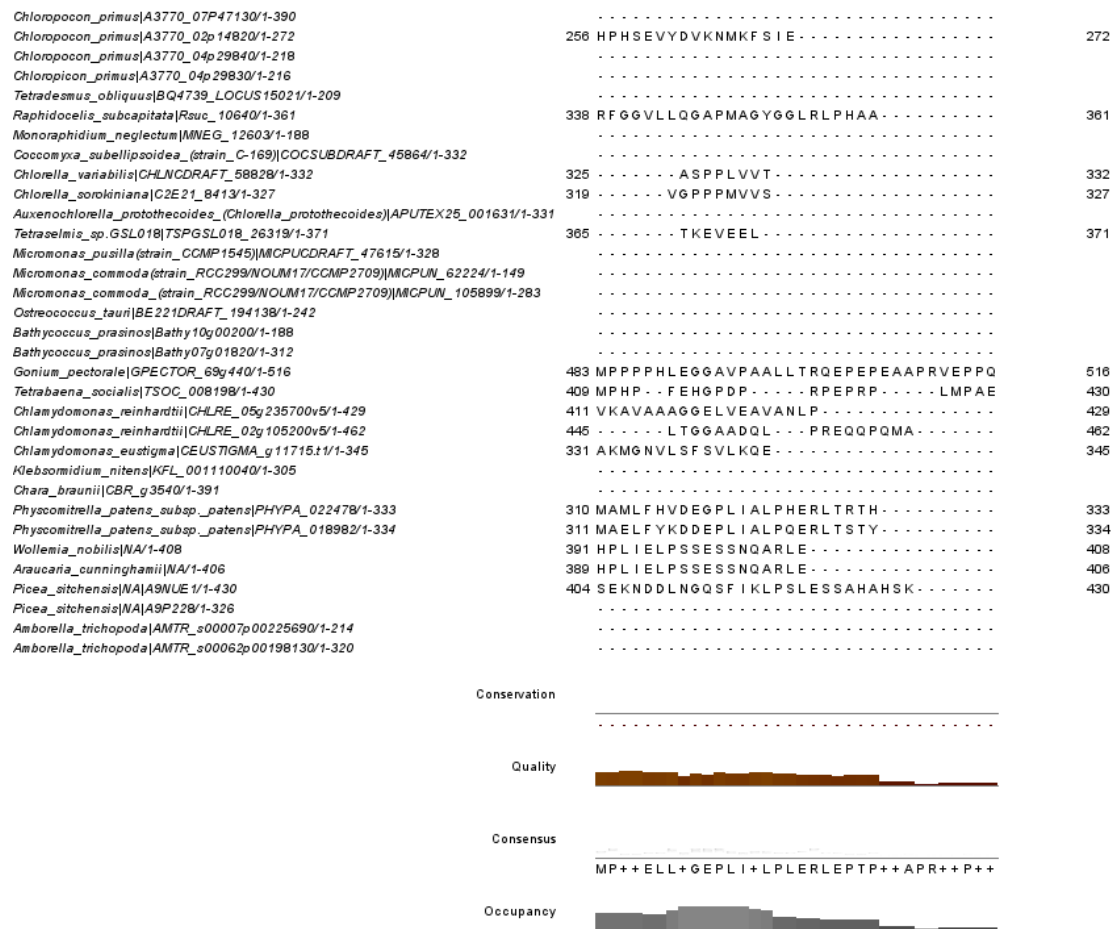


Figure S07. Sequence alignment of PSAPLIPs in green algae, liverwort, moss and gymnosperm. Left: Species names with gene ID after the vertical line. If gene ID was not available, protein is was annotated. Conserved sites were shaded with colors in JalViews. Conservation, quality, consensus, and occupancy were calculated and visualized in JalViews by default.

Panicum_halli_var_halli|GQ65_3G0007700/t-229
Panicum_halli_var_halli|GQ65_3G333500/t-224
Panicum_miliaceum|C2845_PW06504430/t-263
Panicum_miliaceum|C2845_PW17050420/t-231
Panicum_miliaceum|C2845_PW05G21750/t-204
Panicum_miliaceum|C2845_PW07G00670/t-203
Panicum_miliaceum|C2845_PW06G26640/t-184
Setaria_italica|SETIT_7G327400v2/t-230
Setaria_italica|SETI_GG117400v2/t-227
Setaria_italica|SETT_8G015000v2/t-226
Setaria_italica|SETIT_3G284400v2/t-225
Setaria_vindis|SEVIR_7G332200v2/t-230
Setaria_vindis|SEVIR_GG113900v2/t-227
Setaria_vindis|SEVIR_8G013800v2/t-226
Setaria_vindis|SEVIR_3G292600v2/t-225
Aquilegia_scolecias|AQUCO_00400489v1/t-233
Macleaya_cordata|BVCB0_1837g47/t-262
Macleaya_cordata|BVCB0_9017g10/t-212
Papaver_comniferum|C616T_002404/t-357
Nelumbo_nucifera|LOC104597199/t-243
Nelumbo_nucifera|LOC104602464/t-228
Spinacia_oleracea|SOVFE_050110/t-231
Actinidia_chinensis_var_chinensis|CEY00_Acc01859/t-298
Actinidia_chinensis_var_chinensis|CEY00_Acc00072/t-246
Actinidia_chinensis_var_chinensis|CEY00_Acc08725/t-244
Actinidia_chinensis_var_chinensis|CEY00_Acc01859/t-227
Davidia_involucrata|Din_006700/t-269
Davidia_involucrata|Din_026378/t-245
Nyssa_sinenais|F056Z_017856/t-239
Nyssa_sinenais|F056Z_015152/t-125
Atenaria_annua|CTI12_AA282550/t-232
Atenaria_annua|CTI12_AA34949/t-199
Helianthus_annuus|HannXRR_Chrl0g028629/t-181
Cynara_cardunculus_var_oolymus|Cord_003009/t-231
Lactuca_sativa|L_SAT_9X3808/t-229
Daucus_carota_subsp_sativus|DCAR_010960/t-241
Daucus_carota_subsp_sativus|DCAR_018655/t-143
Doronicum_hymetrium|F51_1_2948/t-190
Erythranthe_guttata|MMGLU_ngv1a013247ng/t-226
Geniacea_kurea|JMS69_00799/t-243
Hemdroanthus_impetiginosus|GDL12_11605/t-229
Stipa_asiatica|STAS_21183/t-236
Stipa_asiatica|STAS_234394/t-199
Coffea_campochia|SCOC_T0002323400/t-294
Cuscuta_australis|DM860_002763/t-226
Cuscuta_campestris|CCAM_LOGUS31065/t-226
Cuscuta_campestris|CCAM_LOGUS32789/t-223
Nicotiana attenuata|A4449_38799/t-245
Nicotiana attenuata|A4449_18699/t-238
Nicotiana_sylvestris|LOC104216409/t-245
Nicotiana_sylvestris|LOC104224609/t-238
Nicotiana_tabacum|LOC107812754/t-245
Nicotiana_tabacum|LOC107816607/t-245
Nicotiana_tabacum|LOC107792809/t-238
Nicotiana_tabacum|LOC10777249/t-238
Capsicum_annuum|LOC107843427/t-241
Capsicum_annuum|LOC107851224/t-124
Capsicum_baccatum|CQW23_24170/t-241
Capsicum_baccatum|CQW23_32279/t-197
Capsicum_baccatum|CQW23_29499/t-163
Capsicum_chinense|BC332_26027/t-241
Capsicum_chinense|BC332_31415/t-121
Solanum_chacoense|WAI|A0A0V04M7/t-278
Solanum_chacoense|WAI|A0A0V04W7/t-273
Solanum_tuberosum|102602502/t-242
Solanum_tuberosum|WAI|A0A3030780/t-238
Vitis_vinifera|WIT_08a0058g1030v1-278
Vitis_vinifera|WITSV_040420/t-239
Vitis_vinifera|Paapl_1_v1-195
Vitis_vinifera|WITSV_040421/t-174
Vitis_vinifera|CAK23_0503/t-150
Vitis_vinifera|WAI|Q8861/t-174
Arachis_hypogaea|Ahy_B04g071409/t-245
Arachis_hypogaea|Ahy_A03g006469/t-217
Arachis_hypogaea|Ahy_B03g061501/t-217
Arachis_hypogaea|Ahy_A04g018839/t-212
Lupinus_angustifolius|TarglG_23498/t-222
Lupinus_angustifolius|TarglG_19379/t-204
Cicer_arietinum|LOC101491522/t-279
Cicer_arietinum|LOC101508260/t-215
Medicago_truncatula|MTR_7g072560/t-242
Medicago_truncatula|MtunA11_Chrl6g001314/t-233
Medicago_truncatula|MTR_028040/t-215
Trifolium pratense|L_196_g026334/t-194
Trifolium_subterraneum|TSLUD_266660/t-215
Trifolium_subterraneum|TSLUD_160390/t-181
Lotus_japonicus|WAI|33989/t-216
Cajanus_cajan|KIKL_038939/t-221
Cajanus_cajan|KIKL_029931/t-217
Mucuna_pruriens|CR513_15124/t-288
Mucuna_pruriens|CR513_85238/t-242
Phaseolus_vulgaris|PHAVU_0080084800g/t-222
Phaseolus_vulgaris|PHAVU_0080084700g/t-217
Glycine_max|GLYMA_18G21100/t-256
Glycine_max|GLYMA_19G11400/t-237
Glycine_max|GLYMA_09G277100/t-237
Glycine_max|GLYMA_18G211900/t-236
Glycine_max|GLYMA_01G131400/t-216
Glycine_max|GLYMA_09G277200/t-212
Glycine_max|GLYMA_04G16960/t-202
Glycine_max|GLYMA_19G111500/t-181
Glycine_eqa|DOY65_025469/t-265
Glycine_eqa|DOY65_001399/t-253
Glycine_eqa|DOY65_025469/t-237
Glycine_eqa|DOY65_049180/t-229
Vigna_angulata_var_angulata|Vigan_04G117700/t-252
Vigna_angulata_var_angulata|Vigan_04G117800/t-219
Vigna_angulata_var_angulata|Vigan_09G109000/t-217
Vigna_radiata_var_radiata|LOC106798717/t-219
Vigna_radiata_var_radiata|LOC106754929/t-217
Vigna_radiata_var_radiata|LOC106789949/t-211
Vigna_unguiculata|DEO72_LG10g3244/t-238
Vigna_unguiculata|DEO72_LG10g3245/t-220
Vigna_unguiculata|DEO72_LG8g1152/t-217
Gltus_unshia|GUMV_001140/t-204
Acer_yangbiense|E2V6Z_016741/t-188
Eucalyptus_grandis|EUGRSUZ_K01273/t-227
Eucalyptus_grandis|EUGRSUZ_A00687/t-219
Punica_granatum|CRG98_041613/t-272
Punica_granatum|CRG98_041612/t-227
Punica_granatum|CRG98_016880/t-220
Conthos_villosus|GCOLU_1_2987/t-226
Conthos_villosus|GCOL_30004/t-226
Gossypium_arboreum|F383_27015/t-233
Gossypium_arboreum|F383_21360/t-227
Gossypium_barbadense|GOBAR_AA12144/t-247
Gossypium_barbadense|GOBAR_AA02653/t-247
Gossypium_barbadense|E5319_D10G128500/t-233
Gossypium_barbadense|E5319_A10G160500v1/t-233
Gossypium_barbadense|E5319_D02G005800v1/t-227
Gossypium_barbadense|E5319_A02G004900v1/t-227
Gossypium_darwini|E5288_D10G138900v1/t-233
Gossypium_darwini|E5288_A10G175600v1/t-233
Gossypium_darwini|E5288_A02G005100v1/t-227
Gossypium_darwini|E5288_D02G001700v1/t-227
Gossypium_hirsutum|LOC107896759/t-233
Gossypium_hirsutum|LOC107914554/t-233
Gossypium_hirsutum|LOC107936969/t-227
Gossypium_hirsutum|LOC10793679/t-227
Gossypium_mustelinum|E1491_D10G132600v1/t-233
Gossypium_mustelinum|E1491_A10G165200v1/t-233

Occupancy

175

176

177

[illegible]

203	KAPLF
20T	RVQVI
20S	TTSPA
20D	KTFFF
20D	
2P0	
2KE	
20S	KAPLF
20S	TTSPA
21T	RVQVT
20S	KAPLF
2CP	TTRLA
20S	KAPFF
2	RVQVT
20V	KV-GL
3SLISVKAA	RVFFF
20V	KG-GL
370V	RG-GL
20V	RG-GL
2TV	RVRLF
2DV	RV-GL
200S	LVTDO
20V	KI-GL
20V	PI-GL
20V	RV-GL
20V	RV-GL
20V	RV-GL
20V	RV-GL
1	
200	RI-GL
2VS	
20K	
200	RI-GL
200	KL-GL
20M	RL-EF
2D1	
2DT	RVLLL
2NT	RVLLV
2DT	RVLVF
2DT	RVLVF
53DV	WVPFF
2DM	KACLI
2DM	KA
2DM	KA
20V	RM-GL
2DV	KV-CL
2DF	RVCLC
2DV	KV-CL
2DF	RVCLC
2DF	RV-CL
2DV	KV-CL
2DV	KV-CL
1	
2DV	KV-GL
20S	RL-GL
20S	RLCLA
2DV	KV
2NN	NLEE0
2DL	RL-CL
2DV	KV-CL
20S	RV-GL
2DL	RL-CL
48DV	RV-GL
16SKFKWWE	RAXGL
2CL	PVPAV
2CL	PVPAV
20EG	KM-GL
2EG	RIM0F
2EG	RIM0F
2AE	RM-LL
2AE	RLI0L
300G	RI-GL
2EG	KI-GL
20S	RV-GL
5TL	ERN5L
2EG	KI-0F
20Q	
2EG	KV-GL
200	RI-GL
2EG	RV-GL
2EG	RL-GL
2EG	RV-GL
83EG	RI-GL
290V	RV-GL
2EG	RV-GL
2EG	RMTL
15EG	RM-GL
2EG	RM-GL
2EG	RM-GL
2EG	RT-GL
2NV	EL-0N
2EG	RT-GL
55EG	RT-GL
30EG	RM-GL
2EG	RM-GL
10EG	RM-GL
2EG	RM-AL
2EG	RM-GL
2EG	RM-GL
2EG	RM-GL
2EG	RM-GL
200	RM-AL
2EG	RM-GL
2EG	RM-GL
12AA	OF-NA
30L	RV-GL
20V	KI-0F
2EG	KI-GL
42DV	RI-0F
2AV	RLAFV
2DT	RV-GL
2DT	RV-GL
8DV	RL-GL
2DA	RF-GL
8DV	RV-GL
8DV	RV-GL
8DV	RV-GL
2DA	RF-GL
8DV	RV-GL
2DA	RF-GL
2DA	RF-GL
8DV	RV-GL
8DV	RV-GL

[illegible]

Conservation

Quality

Consensus

Occupancy

00 00 01 01 111 1 0100

GVSF+E+RVGGVLVF+LLLFQSHRHSILLQLNPNLQTHGYTGFMALIEVSLFLFLL+++VLLGA

179

180

182

183

185

Conservation

Quality

Consensus

Occupancy

187

<i>Panicum_hallii_var_hallii</i> [GG65_3G0007700/t-229	114	KEFKQYQVLRDIA	LF	S	130		
<i>Panicum_hallii_var_hallii</i> [GG65_3G333500/t-224	110	EDFVSTSFGEAK	FI	R	126		
<i>Panicum_millicaceum</i> [C2845_PW0000430/t-263	150	EDFVSVHNGNGM	KI	S	166		
<i>Panicum_millicaceum</i> [C2845_PW07G00420/t-231	113	EEFKQYQVLRDIT	LF	SAS	131		
<i>Panicum_millicaceum</i> [C2845_PW05G21750/t-204	90	EDFGISISFGAEAK	LI	R	105		
<i>Panicum_millicaceum</i> [C2845_PW07G00670/t-203	90	EKFRESVHLCKNGM	KI	S	106		
<i>Panicum_millicaceum</i> [C2845_PW06G26640/t-184	70	EDFVSIISFGGEAK	FI	R	86		
<i>Setaria_italica</i> [SETT_7G327400/t-230	114	EEFKQYQVLRDIA	HL	S	130		
<i>Setaria_italica</i> [SETT_8G117400/t-227	114	EDFVSVHNGNGM	KI	S	150		
<i>Setaria_italica</i> [SETT_8G015000/t-226	110	EEFKQYQVLRDIA	LF	S	126		
<i>Setaria_italica</i> [SETT_3G284400/t-225	110	EDFVSVSFGAEAT	FI	R	126		
<i>Setaria_vindis</i> [SEVR_7G337200/t-230	114	EEFKQYQVLRDIA	HL	S	130		
<i>Setaria_vindis</i> [SEVR_8G113900/t-227	114	EKFRESVHLCKNGM	KI	S	130		
<i>Setaria_vindis</i> [SEVR_8G013800/t-226	110	EEFKQYQVLRDIA	LF	S	126		
<i>Setaria_vindis</i> [SEVR_3G283800/t-225	110	EDFVSVSFGAEAT	FI	R	126		
<i>Aquilegia_scoenleai</i> [AQUCO_00400489/t-1-223	100	VEFGSKMNLGIME			122		
<i>Macleaya_cordata</i> [BVC80_1837g47/t-262	175	SEFKKLLNGQTOL	AA	T	161		
<i>Macleaya_cordata</i> [BVC80_9017g10/t-212	101	GDFEKMNLGDE			112		
<i>Papaver_commifera</i> [C6167_002404/t-357	236	ENFHKMDLDAQH			252		
<i>Helumbo_nucleifera</i> [LOC104597159/t-243	129	EEFKVNLGNKML	L	ST	P	145	
<i>Helumbo_nucleifera</i> [LOC10460546/t-1-228	124	LDFFRKMDLGRHRI		IA	S	140	
<i>Spinacia_oleracea</i> [SOIVF_050110/t-231	113	EDFGKKVDLGRNTM	V	SS	M	130	
<i>Actinidia_chinensis_var_chinensis</i> [CEY00_Acc01859/t-258	130	GFLGKGFNFSEHNT		L	MT	S	147
<i>Actinidia_chinensis_var_chinensis</i> [CEY00_Acc00072/t-246	129	GDFRKVNLGGKV		I	S	145	
<i>Actinidia_chinensis_var_chinensis</i> [CEY00_Acc08729/t-244	127	GDLCKVNLGEQV		I	S	143	
<i>Actinidia_chinensis_var_chinensis</i> [CEY00_Acc01859/t-227	90	GFLGKGFNFSEHNT		L	MT	S	147
<i>Davidia_involucrata</i> [Dn_006700/t-269	122	RDFRKVNLFDOMA	IT	S	117		
<i>Davidia_involucrata</i> [Dn_026378/t-245	128	GDFRKVNLGNQMV		I	T	S	144
<i>Nyssa_sinenais</i> [F0562_017856/t-239	128	GDFRKVNLGGMV		I	T	S	144
<i>Nyssa_sinenais</i> [F0562_015152/t-125	26	GDFCKVNLGGMA		I	P	S	46
<i>Athenia_annua</i> [C712_AA362595/t-192	121	AEEFKVNLGGQV		AY	A	137	
<i>Athenia_annua</i> [C712_AA349490/t-199	86	AEEFKVNLGGQV		AY	A	104	
<i>Helianthus_annuus</i> [HannXRR_Chrl0g028629/t-1-181	70	EDFCKVNLCKEUV		AY	A	86	
<i>Cynara_cardunculus_var_cylindrus</i> [Cord_003008/t-231	120	SDFCKVNLGNEV		AY	A	136	
<i>Lactuca_satiava</i> [SAT_3X3808/t-1-229	121	EDFCKVNLCKEIV		F	VS	A	137
<i>Daucus_carota_subsp_sativus</i> [DCAR_010960/t-241	124	ADFKVNLGQV		F	VS	E	141
<i>Daucus_carota_subsp_sativus</i> [DCAR_010655/t-143	30	EDFCEMDEKQIA		LI	A	S	54
<i>Daucocera_hymenotum</i> [F51_29469/t-190	73	DGDFCKVNLGEEVI		S	FP	L	60
<i>Erythranthe_guttata</i> [MMIGU_Lngv1a013247g/t-226	115	EDFLKVDLGEKKA		SV	A	131	
<i>Genlisea_aurea</i> [M569_00799/t-243	119	GEFFKQGLGEGAI			SS		134
<i>Hedroanthus_pegelinoaus</i> [COL12_11605/t-229	112	EDFRKVDLGEQV		SI	S	128	
<i>Stipa_saxatilis</i> [STAS_21163/t-236	116	DALCKVNLGEEG		ST	S	134	
<i>Stipa_saxatilis</i> [STAS_23420/t-199	95	EDFRKVDLGEQV		ST	S	134	
<i>Coffea_caneholia</i> [GSCOG_7000232400/t-1-294	176	EDFKVDLGEDIV		SI	S	162	
<i>Cuscuta_australis</i> [DM860_002763/t-226	116	EQLCKKAAMCKVT		L	T	S	132
<i>Cuscuta_campestris</i> [CCAM_LOCUS31085/t-226	116	EDFKKAAMCKVT		L	T	S	132
<i>Cuscuta_campestris</i> [CCAM_LOCUS32789/t-223	113	EDFKKAAMCKVT		L	T	S	129
<i>Nicotiana_glabra</i> [NAG_02789/t-245	125	EDFEKFDLGERVV		SI	S	141	
<i>Nicotiana attenuata</i> [A4A4_19559/t-238	121	DDI CKVDLCKV		SI	S	137	
<i>Nicotiana_sylvestris</i> [LOC104216406/t-245	125	EDFEKFDLGERVV		TI	S	141	
<i>Nicotiana_sylvestris</i> [LOC104224609/t-238	121	DDI CKVDLCKV		SI	S	137	
<i>Nicotiana_sylvestris</i> [LOC107812754/t-245	125	EDFEKFDLGERVV		TV	S	141	
<i>Nicotiana_sylvestris</i> [LOC107816607/t-245	125	EDFEKFDLGERVV		TI	S	141	
<i>Nicotiana_sylvestris</i> [LOC107826091/t-238	121	DDI CKVDLCKV		SI	S	137	
<i>Nicotiana_sylvestris</i> [LOC10777349/t-238	121	DDI CKVDLCKV		SI	S	137	
<i>Capiscum_annuum</i> [LOC10784342/t-1-241	124	DDI CKIHLCKV		SI	S	140	
<i>Capiscum_annuum</i> [LOC10785124/t-1-124	91	EEFEKFDLCKV		L	O	103	
<i>Capiscum_baccatum</i> [COW33_24170/t-241	124	DDI CKIHLCKV		SI	S	140	
<i>Capiscum_baccatum</i> [COW33_32779/t-197	166	EEFF...GFDLCKV		L	O	176	
<i>Capiscum_baccatum</i> [COW33_29495/t-163	131	EEFEKFDLCKV		L	O	140	
<i>Capiscum_chinense</i> [B332_26027/t-241	124	DDI CKIHLCKV		SI	S	140	
<i>Capiscum_chinense</i> [B332_31415/t-121	88	EEFEKFDLCKV		L	O	100	
<i>Solanum_chacoense</i> [WAI4A040V04M7/t-278	125	ENFGQDFLGEQV		I	IS	Q	142
<i>Solanum_chacoense</i> [WAI4A040V04M7/t-273	124	DDI CKVDLCKV		SI	S	Q	141
<i>Solanum_tuberosum</i> [LOC10782609/t-238	125	ENFGQDFLGEQV		I	IS	Q	142
<i>Solanum_tuberosum</i> [WAI4A040V04M7/t-278	125	ENFGQDFLGEQV		I	IS	Q	142
<i>Vitis_vinifera</i> [VT_08c0058/t-1030/t-278	172	GNFGQDFLGEQV		TS	P	188	
<i>Vitis_vinifera</i> [VITSV_040420/t-239	133	GNFGQDFLGEQV		TS	P	146	
<i>Vitis_vinifera</i> [Paapl_1/t-195	89	GNFGQDFLGEQV		TS	P	105	
<i>Vitis_vinifera</i> [VITSV_040420/t-174	61	GNFEKANLGPST		TA	G	77	
<i>Vitis_vinifera</i> [C203_030315/t-150	89	GNFGQDFLGEQV		TS	P	104	
<i>Vitis_vinifera</i> [WAI4A040V04M7/t-278	61	GNFEKANLGPST		TA	G	77	
<i>Arachis_hypogaea</i> [Ahy_804g071408/t-245	128	EDFKVNLGERT		KI	S	144	
<i>Arachis_hypogaea</i> [Ahy_804g071408/t-245	100	QELSTANIPFILL		NV		124	
<i>Arachis_hypogaea</i> [Ahy_804g071408/t-245	100	QELSTANIPFILL		NV		124	
<i>Arachis_hypogaea</i> [Ahy_804g071408/t-245	100	QELSTANIPFILL		NV		124	
<i>Lupinus_angustifolius</i> [Tanji_22468/t-222	113	EEFKVNLGERT		KI	S	118	
<i>Lupinus_angustifolius</i> [Tanji_22468/t-222	113	EEFKVNLGERT		KI	S	118	
<i>Cicer_arietinum</i> [LOC101491523/t-279	111	RELCKVNLGPYSA		KI	S	127	
<i>Cicer_arietinum</i> [LOC101508260/t-215	146	ODFKVNLGOTIA		D	LS	L	165
<i>Medicago_frunctosa</i> [MTR_7072560/t-242	98	EELCKVNLGESA		AS	S	114	
<i>Medicago_frunctosa</i> [MTR_7072560/t-242	100	GDFCKVNLGONIA		N	I	S	125
<i>Medicago_frunctosa</i> [MTR_7072560/t-242	100	AEELCKVNLGESA		IS	T	125	
<i>Medicago_frunctosa</i> [MTR_7072560/t-242	98	AEELCKVNLGESA		IS	T	125	
<i>Trifolium pratense</i> [L195_g026334/t-194	77	EELCKVNLGESA		NY	A	93	
<i>Trifolium pratense</i> [L195_g026334/t-194	98	EELCKVNLGESA		NY	A	114	
<i>Trifolium pratense</i> [L195_g026334/t-194	112	GDFCKVNLGONIA		N	F	S	128
<i>Lotus japonicus</i> [WAI33989/t-216	96	EELCKVNLGESA		LS	S	114	
<i>Cajanus cajan</i> [KCI_039931/t-221	110	EDFKVNLGONIA		Y	I	S	116
<i>Cajanus cajan</i> [KCI_039931/t-221	100	EELCKVNLGONIA		V	S	S	116
<i>Mucuna pruriens</i> [CR513_15124/t-288	179	EDFKVNLGONIA		Y	V	S	106
<i>Mucuna pruriens</i> [CR513_15124/t-288	125	EELCKVNLGONIA		V	S	S	141
<i>Phaseolus vulgaris</i> [PHAVI_008004800/t-1-222	111	REFCHKIDILQTE		HI	S	127	
<i>Phaseolus vulgaris</i> [PHAVI_008004800/t-1-222	102	EDFKVNLGONIA		Y	I	S	116
<i>Glycine_max</i> [SLYMA_19G212100/t-250	115	EDFKVNLGOLIT		YI	S	131	
<i>Glycine_max</i> [SLYMA_19G212100/t-250	102	EDFKVNLGOLIT		YI	S	118	
<i>Glycine_max</i> [SLYMA_19G212100/t-250	102	EDFKVNLGOLIT		YI	S	118	
<i>Glycine_max</i> [SLYMA_19G212100/t-250	125	EDFKVNLGODIA		HI	S	141	
<i>Glycine_max</i> [SLYMA_19G212100/t-250	96	EELCKVNLGOLIT		KI	S	116	
<i>Glycine_max</i> [SLYMA_19G212100/t-250	102	EDFKVNLGOLIT		YI	S	118	
<i>Glycine_max</i> [SLYMA_19G212100/t-250	86	EELCKVNLGOLIT		YI	P	101	
<i>Glycine_max</i> [SLYMA_19G212100/t-250	101	EDFKVNLGONIA		YI	G	117	
<i>Glycine_max</i> [SLYMA_19G212100/t-250	154	EDFKVNLGODIA		YI	A	170	
<i>Glycine_max</i> [SLYMA_19G212100/t-250	136	EELCKVNLGOLIT		KI	S	162	
<i>Glycine_max</i> [SLYMA_19G212100/t-250	102	EDFKVNLGOLIT		YI	S	118	
<i>Glycine_max</i> [SLYMA_19G212100/t-250	119	EDFKVNLGODIA		HI	S	154	
<i>Vigna_angularis_var_angularis</i> [Vigan_04G117700/t-252	116	EDFKVNLGOLIT		YI	S	133	
<i>Vigna_angularis_var_angularis</i> [Vigan_04G117700/t-252	108	REFCHKIDILQTE		HV	S	124	
<i>Vigna_angularis_var_angularis</i> [Vigan_04G117700/t-252	99	RELCKVNLGONIA		LI	S	115	
<i>Vigna_angularis_var_angularis</i> [Vigan_04G117700/t-252	108	REFCHKIDILQTE		HV	S	124	
<i>Vigna_angularis_var_angularis</i> [Vigan_04G117700/t-252	99	RELCKVNLGONIA		LI	S	115	
<i>Vigna_angularis_var_angularis</i> [Vigan_04G117700/t-252	108	REFCHKIDILQTE		HV	S	124	
<i>Vigna_angularis_var_angularis</i> [Vigan_04G117700/t-252	99	RELCKVNLGONIA		LI	S	115	
<i>Vigna_angularis_var_angularis</i> [Vigan_04G117700/t-252	108	REFCHKIDILQTE		HV	S	124	
<i>Vigna_angularis_var_angularis</i> [Vigan_04G117700/t-252	99	RELCKVNLGONIA		LI	S	115	
<i>Vigna_angularis_var_angularis</i> [Vigan_04G117700/t-252	108	REFCHKIDILQTE		HV	S	124	
<i>Vigna_angularis_var_angularis</i> [Vigan_04G117700/t-252	99	RELCKVNLGONIA		LI	S	115	
<i>Vigna_angularis_var_angularis</i> [Vigan_04G117700/t-252	108	REFCHKIDILQTE		HV	S	124	
<i>Vigna_angularis_var_angularis</i> [Vigan_04G117700/t-252	99	RELCKVNLGONIA		LI	S	115	
<i>Vigna_angularis_var_angularis</i> [Vigan_04G117700/t-252	108	REFCHKIDILQTE		HV	S	124	
<i>Vigna_angularis_var_angularis</i> [Vigan_04G117700/t-252	99	RELCKVNLGONIA		LI	S	115	
<i>Vigna_angularis_var_angularis</i> [Vigan_04G117700/t-252	108	REFCHKIDILQTE		HV	S	124	
<i>Vigna_angularis_var_angularis</i> [Vigan_04G117700/t-252	99	RELCKVNLGONIA		LI	S	115	
<i>Vigna_angularis_var_angularis</i> [Vigan_04G117700/t-252	108	REFCHKIDILQTE		HV	S	124	
<i>Vigna_angularis_var_angularis</i> [Vigan_04G117700/t-252	99	RELCKVNLGONIA		LI	S	115	
<i>Vigna_angularis_var_angularis</i> [Vigan_04G117700/t-252	108	REFCHKIDILQTE		HV	S	124	
<i>Vigna_angularis_var_angularis</i> [Vigan_04G117700/t-252	99	RELCKVNLGONIA		LI	S	115	
<i>Vigna_angularis_var_angularis</i> [Vigan_04G117700/t-252	108	REFCHKIDILQTE		HV	S	124	
<i>Vigna_angularis_var_angularis</i> [Vigan_04G117700/t-252	99	RELCKVNLGONIA		LI	S	115	
<i>Vigna_angularis_var_angularis</i> [Vigan_04G117700/t-252	108	REFCHKIDILQTE		HV	S	124	
<i>Vigna_angularis_var_angularis</i> [Vigan_04G117700/t-252	99	RELCKVNLGONIA		LI	S	115	
<i>Vigna_angularis_var_angularis</i> [Vigan_04G117700/t-252	108	REFCHKIDILQTE		HV	S	124	
<i>Vigna_angularis_var_angularis</i> [Vigan_04G117700/t-252	99	RELCKVNLGONIA		LI	S	115	
<i>Vigna_angularis_var_angularis</i> [Vigan_04G117700/t-252	108	REFCHKIDILQTE		HV	S	124	
<i>Vigna_angularis_var_angularis</i> [Vigan_04G117700/t-252	99	RELCKVNLGONIA		LI	S	115	
<i>Vigna_angularis_var_angularis</i> [Vigan_04G117700/t-252	108	REFCHKIDILQTE		HV	S	124	
<i>Vigna_angularis_var_angularis</i> [Vigan_04G117700/t-252	99	RELCKVNLGONIA		LI	S	115	
<i>Vigna_angularis_var_angularis</i> [Vigan_04G117700/t-252	108	REFCHKIDILQTE		HV	S	124	
<i>Vigna_angularis_var_angularis</i> [Vigan_04G117700/t-252	99	RELCKVNLGONIA		LI	S	115	
<i>Vigna_angularis_var_angularis</i> [Vigan_04G117700/t-252	108	REFCHKIDILQTE		HV	S	124	
<i>Vigna_angularis_var_angularis</i> [Vigan_04G117700/t-252	99	RELCKVNLGONIA		LI	S	115	
<i>Vigna_angularis_var_angularis</i> [Vigan_04G117700/t-252	108	REFCHKIDILQTE		HV	S	124	
<i>Vigna_angularis_var_angularis</i> [Vigan_04G117700/t-252	99	RELCKVNLGONIA		LI	S	115	
<i>Vigna_angularis_var_angularis</i> [Vigan_04G117700/t-252	108	REFCHKIDILQTE		HV	S	124	
<i>Vigna_angularis_var_angularis</i> [Vigan_04G117700/t-252	99	RELCKVNLGONIA		LI	S	115	
<i>Vigna_angularis_var_angularis</i> [Vigan_04G117700/t-252	108	REFCHKIDILQTE		HV	S	124	
<i>Vigna_angularis_var_angularis</i> [Vigan_04G117700/t-252	99	RELCKVNLGONIA		LI	S	115	
<i>Vigna_angularis_var_angularis</i> [Vigan_04G117700/t-252	108	REFCHKIDILQTE		HV	S	124	
<i>Vigna_angularis_var_angularis</i> [Vigan_04G117700/t-252	99	RELCKVNLGONIA		LI	S	115	
<i>Vigna_angularis_var_angularis</i> [Vigan_04G117700/t-252	108	REFCHKIDILQTE		HV	S	124	
<i>Vigna_angularis_var_angularis</i> [Vigan_04G117700/t-252	99	RELCKVNLGONIA		LI	S	115	
<i>Vigna_angularis_var_angularis</i> [Vigan_04G117700/t-252	108	REFCHKIDILQTE		HV	S	124	
<i>Vigna_angularis_var_angularis</i> [Vigan_04G117700/t-252	99	RELCKVNLGONIA		LI	S	115</	

Conservation

228 245 28 1143 00 00

Quality

Consensus

5' FC - - - - - LC

EEFCKKVKLQCKVA++IPLSTSS++SLFSPDDN+LLLS+SS+LEV+*****K+*****L

Occupancy





189

190

<i>Vitis_rotundifolia</i> WJQ06614/1-174	78Q-L-Y-Q-D.....	S.....	V V F R R M L V M K V L I K L R D P D T.....	103
<i>Arachis_hypogaea</i> Ahy_B04q071408/1-245	145L-Q-I-Q-E.....	N-S.....	..D F E D A M S A L L A K L K D P D T.....	171
<i>Arachis_hypogaea</i> Ahy_A03q006489/1-217	125-Q-E.....	D-S.....	..S A K R A T V S A I L V K L K D P E T.....	148
<i>Arachis_hypogaea</i> Ahy_B03q06150/1-217	125-Q-Q.....	D-S.....	..S A K R A T V S A I L I K L K D P E T.....	148
<i>Arachis_hypogaea</i> Ahy_A04q0178839/1-212	119-Q-I-Q-E.....	N-S.....	..D F E D A M S A L L A K L K D P D T.....	144
<i>Lupinus_angustifolius</i> TanjiG_23489/1-222	131-Q-A-E-O.....	N-S.....	..D F R K D T V S A L L L K L N D P D T.....	156
<i>Lupinus_angustifolius</i> TanjiG_19378/1-204	128-S-Q-V-Q-E.....	N-N.....	..D L Q K D T I T S L V A K L K D P D T.....	154
<i>Cicer_arietinum</i> LOC101491522/1-279	166-Q-V-Q-E.....	N-S.....	..R E F E E D T V S A L L A K I K D P D T.....	181
<i>Cicer_arietinum</i> LOC101508260/1-215	115-Q-V-Q-G-N.....	S.....	..G O L C K D A V A A L L V E N D P D T.....	140
<i>Medicago_truncatula</i> WTR_072560/1-242	126L-K-V-Q-E.....	N-T.....	..R E F E E T T V S S L L D K I K D P D T.....	153
<i>Medicago_truncatula</i> WTR_02940/1-215	123-Q-V-Q-E.....	N-S.....	..R E F R K D T V A E L L V E N D P D T.....	145
<i>Medicago_truncatula</i> WTR_02904/1-215	115-Q-V-H-G-N.....	S.....	..D F R K D T V A E L L V E N D P E T K.....	141
<i>Trifolium pratense</i> LT195_q026334/1-194	94-R-V-R-G-N.....	S.....	..D F R K D T V A Q L L V E N D P D T.....	119
<i>Trifolium_subterraneum</i> ITSU02_066660/1-215	115-R-V-G-G-G.....	N-S.....	..D F R K D T V A Q L L V E N D P D T K.....	142
<i>Trifolium_subterraneum</i> ITSU02_160380/1-181	129L-Q-V-Q-E.....	N-S.....	..R E F R K D T I T S S L L D K I K D P D T.....	155
<i>Lobelia japonica</i> WJQ35989/1-216	115-Q-V-Q-E.....	S.....	..D F R K D A V S A L L V E N D P D T.....	140
<i>Cajanus cajan</i> KKK_L_035920/1-221	128-Q-V-Q-E.....	D-S.....	..R E F R D D T V S T L L A K L K D P D T.....	153
<i>Cajanus cajan</i> KKK_L_02993/1-217	117-Q-V-K-G-N.....	S.....	..D S S K D S V A A L L V K L N D P D T.....	142
<i>Mucuna pruriens</i> CR513_15124/1-288	197-K-D-Q-E.....	D-T.....	..D F R K D T V S T L L E K L E S D I.....	222
<i>Mucuna pruriens</i> CR513_55238/1-242	142-Q-V-Q-G-N.....	T.....	..D S S K D T V L L L A K L S D P D T.....	167
<i>Phaseolus vulgaris</i> PHAVU_0080084800/1-222	128L-Q-V-Q-E.....	D-A.....	..R E F E E T T V S T L L V K L K S D T.....	154
<i>Phaseolus vulgaris</i> PHAVU_0080084700/1-217	119-S-Q-V-Q-E.....	D-S.....	..D F R K D T V S T L L A K L K S D T.....	145
<i>Glycine_max</i> [SLYMA_18G212100/1-250	132L-Q-V-Q-E.....	D-T.....	..S O F R E E D T V S T L L A K L K D P D T.....	158
<i>Glycine_max</i> [SLYMA_19G111400/1-237	119L-L-V-Q-E.....	D-T.....	..S O F R K V T V S T L L A K L K S D T.....	145
<i>Glycine_max</i> [SLYMA_09G277100/1-237	119L-L-V-Q-E.....	D-T.....	..S O F R K D T V S T L S A K L K S D T.....	145
<i>Glycine_max</i> [SLYMA_18G211900/1-236	142L-K-V-Q-E.....	D-S.....	..R E F R K D T V S T L L E K L K E S D T.....	168
<i>Glycine_max</i> [SLYMA_01G131400/1-216	116-E-V--E-O.....	N-S.....	..D S S K D T V S A L L V K L N D P D T.....	141
<i>Glycine_max</i> [SLYMA_09G277200/1-217	119L-K-V-Q-D.....	D-S.....	..R E F R D D T V S T L L E K L K E S D T.....	145
<i>Glycine_max</i> [SLYMA_04G159500/1-202	102L-Q-V-Q-E.....	D-T.....	..S O F R E D I.....	115
<i>Glycine_max</i> [SLYMA_19G111500/1-181	118L-K-V-Q-D.....	D-S.....	..R E F R D D T V S T L L E K L K E T D T.....	144
<i>Glycine_ega</i> [DDB6_025469/1-265	171L-K-V-Q-D.....	D-S.....	..R E F R D D T V S T L L E K L K E S D T.....	197
<i>Glycine_ega</i> [DDB6_001396/1-253	153-E-V--E-G.....	N-S.....	..D S S K D T V S A L L V K L N D P D T.....	178
<i>Glycine_ega</i> [DDB6_025469/1-237	119L-L-V-Q-E.....	D-T.....	..D F R K D T V S T L S A K L K S D T.....	145
<i>Glycine_ega</i> [DDB6_049180/1-229	139L-K-V-Q-E.....	D-S.....	..R E F R K D T V S T L L E K L K E S D T.....	161
<i>Vigna_angularis</i> var_angularis[Vigan_04G117700/1-262	134-Q-V-Q-E.....	D-S.....	..D O F F K D A V S T L L T L K L K S D T.....	159
<i>Vigna_angularis</i> var_angularis[Vigan_04G117800/1-219	125L-Q-V-Q-E.....	D-A.....	..D F R E E T T V S T L L D K L K S D T.....	151
<i>Vigna_angularis</i> var_angularis[Vigan_09G10900/1-217	116-Q-V-Q-G-N.....	K-A.....	..D S S K D T V S A I L V K L N D P D T.....	141
<i>Vigna_radiata</i> var_radiata[LOC106786177/1-219	124S-L-K-V-Q-E.....	D-A.....	..D F R E E T T V S T L L D K L K S D T.....	151
<i>Vigna_radiata</i> var_radiata[LOC106786293/1-217	116-Q-V-Q-G-N.....	K-C.....	..D S S K D T V S A I L V K L N D P E T.....	141
<i>Vigna_radiata</i> var_radiata[LOC106786949/1-211	119-S-Q-V-Q-E.....	D-S.....	..D O F F K D V V S T L L T L K L K S D T.....	145
<i>Vigna_unguiculata</i> DE072_LG10g3244/1-238	120H-V--Q-E.....	D-S.....	..D F R D N A V S T L L A K L K S D T.....	145
<i>Vigna_unguiculata</i> DE072_LG10g3245/1-220	126-V-Q-V-Q-E.....	D-A.....	..D F R E E T T V S T L L D K L K S D T.....	152
<i>Vigna_unguiculata</i> DE072_LG8y1155/1-217	116-Q-V-Q-G-N.....	T.....	..D S S K D T V S A I L V K L N D P D T.....	141
<i>Onosmodium</i> [CJHMF_001141/1-204	150-Q-L--R-E.....	D-S.....	..E L L H H T V S E V L T K L K D P D T.....	175
<i>Acetysanthus</i> [E263_01674/1-186	139-Q-L--Q-D.....	D-S.....	..D O V H H A V S E V L T K L K D P D T.....	163
<i>Eucalyptus_grandis</i> EUGRSUZ_K01273/1-227	127-Q-L--K-E.....	D-S.....	..R E F R D A V S O V L D K L K D P D T.....	152
<i>Eucalyptus_grandis</i> EUGRSUZ_A00887/1-219	119-S-Q-L-Q-E.....	D-S.....	..R E L D H H A V S E V L D K L K D P D T.....	145
<i>Punica granatum</i> CRG98_041613/1-272	171-Q-I--K-E.....	D-S.....	..D F R D N A V S E L M D K L K D P D T.....	196
<i>Punica granatum</i> CRG98_041612/1-227	126-Q-I--K-E.....	D-S.....	..R E F R D N T V S E L M D K L K D P D T.....	151
<i>Punica granatum</i> CRG98_041608/1-230	137-Q-I--R-E.....	D-S.....	..D O V H H A V S E V L T K L K D P D T.....	163
<i>Conthosia capsulata</i> [CCCVL_1_28877/1-226	137-Q-I--R-E.....	D-S.....	..D O L D N A V S O V L E K L K D P D T.....	162
<i>Conthosia dillitoria</i> COLO4_30004/1-226	137-Q-I--R-E.....	D-S.....	..D O L D H N A V S O V L E K L N D P D T.....	162
<i>Gossypium arboreum</i> F383_27015/1-233	144-Q-I--R-E.....	D-C.....	..D O V D H H A V S E V L T K L K D P D T.....	169
<i>Gossypium arboreum</i> F383_21360/1-227	138-Q-F--R-E.....	D-S.....	..D G M D H R A I S E V L M K L D P D T.....	163
<i>Gossypium barbadense</i> [GGBAR_AA12144/1-247	143-S-Q-F--R-E.....	D-S.....	..D G M D H R A I S E V L M K L D P D T K L L D P D V L.....	178
<i>Gossypium barbadense</i> [GGBAR_AA02863/1-247	139-Q-I--R-E.....	D-S.....	..D O V D H H A V S E V L T K L K D P D T.....	163
<i>Gossypium barbadense</i> [E5319_D10G128500/1-1-233	144-Q-I--R-E.....	D-C.....	..D O V D H H A V S E V L T K L K D P D T.....	169
<i>Gossypium barbadense</i> [E5319_A10G160500/1-1-233	144-Q-I--R-E.....	D-C.....	..D O V D H H A V S E V L T K L K D P D T.....	169
<i>Gossypium barbadense</i> [E5319_D02G005800/1-1-227	138-Q-F--R-E.....	D-S.....	..D G M D H R A I S E V L M K L D P D T.....	163
<i>Gossypium barbadense</i> [E5319_A02G004900/1-1-227	138-Q-F--R-E.....	D-S.....	..D G M D H R A I S E V L M K L D P D T.....	163
<i>Gossypium darwini</i> [E5288_D10G138000/1-1-233	144-Q-I--R-E.....	D-C.....	..D O V D H H A V S E V L T K L K D P D T.....	169
<i>Gossypium darwini</i> [E5288_A10G179500/1-1-233	144-Q-I--R-E.....	D-C.....	..D O V D H H A V S E V L T K L K D P D T.....	169
<i>Gossypium darwini</i> [E5288_A02G005100/1-1-227	138-Q-F--R-E.....	D-S.....	..D G M D H R A I S E V L M K L D P D T.....	163
<i>Gossypium darwini</i> [E5288_D02G001700/1-1-227	138-Q-F--R-E.....	D-S.....	..D G M D H R A I S E V L M K L D P D T.....	163
<i>Gossypium hirsutum</i> [LOC10786756/1-233	144-Q-I--R-E.....	D-C.....	..D G V R H H P V S E V L T K L K D P D T.....	169
<i>Gossypium hirsutum</i> [LOC107914554/1-233	144-Q-I--R-E.....	D-C.....	..D O V D H H A V S E V L T K L K D P D T.....	169
<i>Gossypium hirsutum</i> [LOC10793569/1-217	138-Q-F--R-E.....	D-S.....	..D G M D H R A I S E V L M K L D P D T.....	163
<i>Gossypium hirsutum</i> [LOC107903579/1-227	138-Q-F--R-E.....	D-S.....	..D G M D H R A I S E V L M K L D P D T.....	163
<i>Gossypium mustelinum</i> [E1491_D10G132600/1-1-233	144-Q-I--R-E.....	D-C.....	..D O V D H H A V S E V L T K L K D P D T.....	169
<i>Gossypium mustelinum</i> [E1491_A10G165200/1-1-233	144-Q-I--R-E.....	D-C.....	..D O V D H H A V S E V L T K L K D P D T.....	169
<i>Gossypium mustelinum</i> [E1491_D02G005900/1-1-227	138-Q-F--R-E.....	D-S.....	..D G M D H R A I S E V L M K L D P D T.....	163
<i>Gossypium mustelinum</i> [E1491_A02G005100/1-1-227	138-Q-F--R-E.....	D-S.....	..D G M D H R A I S E V L M K L D P D T.....	163
<i>Gossypium raimondii</i> B456_01G112940/1-233	144-Q-I--R-E.....	D-C.....	..D O V D H H A V S E V L T K L K D P D T.....	169
<i>Gossypium raimondii</i> B456_00G050800/1-227	138-Q-F--R-E.....	D-S.....	..D G M D H R A I S E V L M K L D P D T.....	163
<i>Gossypium tomentosum</i> [E5332_D10G139500/1-1-233	144-Q-I--R-E.....	D-C.....	..D O V D H H A V S E V L T K L K D P D T.....	169
<i>Gossypium tomentosum</i> [E5332_A10G178500/1-233	144-Q-I--R-E.....	D-C.....	..D O V D H H A V S E V L T K L K D P D T.....	169
<i>Gossypium tomentosum</i> [E5332_A02G005200/1-1-227	138-Q-F--R-E.....	D-S.....	..D G M D H R A I S E V L M K L D P D T.....	163
<i>Gossypium tomentosum</i> [E5332_D02G005900/1-1-227	138-Q-F--R-E.....	D-S.....	..D G M D H R A I S E V L M K L D P D T.....	163
<i>Theobroma cacao</i> TCM_019744/1-228	139-Q-I--R-E.....	D-S.....	..D G M D H R A I A E V L I K L K D P D T Q.....	164
<i>Arabidopsis thaliana</i> ALAP_AA5G141700/1-213	119-E-A-R-Q.....	D-S.....	..D G V D H R T V S E I I K L K D P D V.....	145
<i>Arabidopsis thaliana</i> ANL_LOCUS23250/1-225	127-S-E-V-H-Q.....	N.....	..R E T R E T V S E V I K L K D P E T.....	153
<i>Arabidopsis thaliana</i> ANL_LOCUS15820/1-214	119-E-A-R-Q.....	D-S.....	..D G V D H R T V S E I I K L K D P D V.....	145
<i>Arabidopsis thaliana</i> ANL_LOCUS15790/1-214	119-E-A-R-Q.....	D-S.....	..D G V D H R T V S E I I K L K D P D V.....	145
<i>Brassicica napus</i> subsp_pinnatifida[WJQ40030313/1-215	120-K-E-A-R-Q.....	D-S.....	..D G V D H R T V S E I I K L K D P D T.....	145
<i>Brassicica napus</i> subsp_pinnatifida[WJQ40030313/1-215	120-E-A-R-Q.....	D-S.....	..D A V D H R T V S D I L I K L K D P D T.....	140
<i>Brassicica napus</i> [BnaC07g32480D/1-216	120-Q-E-A-R-Q.....	D-T.....	..D V D H R T V S E I I K L K D P D T.....	145
<i>Brassicica napus</i> [BnaA03g57960D/1-215	119-E-A-R-Q.....	D-T.....	..D G V D H R T V S E I I K L K D P D T.....	145
<i>Brassicica napus</i> [BnaCmg04660D/1-199	106-Q-G.....	N.....	..D A D H D T V S O L L P K L K D P D T.....	131
<i>Brassicica oleracea</i> var_oleracea[WJQ40030303/1-219	136-E-A-R-Q.....	D-S.....	..D A V D H R T V S D I L I K L K D P D T.....	145
<i>Brassicica oleracea</i> var_oleracea[WJQ40030303/1-216	120-Q-E-A-R-Q.....	D-S.....	..D V D H R T V S E I I K L K D P D T.....	145
<i>Brassicica oleracea</i> var_oleracea[WJQ40030303/1-199	109-Q-G.....	N.....	..D A D H D T V S O L L P K L K D P D T.....	131
<i>Arabidopsis thaliana</i> subsp_lyrata[ARALYDRAFT_486889/1-220	123-Q-V-H-Q.....	N.....	..R E A R D T V S E V V A K L K D P D T.....	148
<i>Arabidopsis thaliana</i> subsp_lyrata[ARALYDRAFT_66800/1-213	119-E-E-V-R-Q.....	D-S.....	..D G V D H R T V S E I I K L K D P D T.....	145
<i>Arabidopsis thaliana</i> [Atg5g01800/1-217	120-Q-V-H-Q.....	N.....	..R E A R D T V S E V V T K L K D P E T.....	145
<i>Arabidopsis thaliana</i> [Atg5g3130/1-213	119-E-A-R-Q.....	D-S.....	..D G V D H R T V S E I I K L K D P D T.....	145
<i>Caprilla rubellia</i> [CARUB_v10001904q/1-223	122-Q-V-H-Q-E.....	N.....	..D A R D R O T V S E V V A K L K D P E T.....	147
<i>Caprilla rubellia</i> [CARUB_v100019623q/1-233	119-E-D-V-R-Q.....	D-S.....	..D G V D H R T V S E I I K L K D P D T.....	145
<i>Eutrema halophilum</i> [WJQ40030313/1-213	119-E-A-R-Q.....	D-S.....	..D G V D H R T V S E I I K L K D P D T Q.....	146
<i>Eutrema halophilum</i> [EUTSA_v10010716q/1-213	119-E-A-R-Q.....	D-S.....	..D G V D H R T V S E I I K L K D P D T.....	145
<i>Eutrema halophilum</i> [EUTSA_v10014673q/1-209	113-P-L-S-H-Q.....	G-N.....	..R E S R O T V S O V V S N K L K D P E I.....	139
<i>Noccaea caenulensis</i> [L_T12690_c0_g1_l1_g_41269/1-219	121L-A-A-S-Q.....	G-N.....	..R E A R D T V T E L V T O L K D P E T.....	146
<i>Noccaea caenulensis</i> [L_T14311_c0_g1_l1_g_15690/1-218	120L-A-A-S-Q.....	G-N.....	..R E A R D T V T E L V T O L K D P E T.....	146
<i>Noccaea caenulensis</i> [GA_T12421_c0_g1_l1_g_3980/1-218	120L-A-A-S-Q.....	G-N.....	..R E A R D T V T E L V T O L K D P E T.....	146
<i>Noccaea caenulensis</i> [MP_T19898_c0_g1_l1_g_2735/1-218	120L-A-A-S-Q.....	G-N.....	..R E A R D T V T E L V T O L K D P E T.....	146
<i>Noccaea caenulensis</i> [MP_T19898_c0_g1_l1_g_44534/1-213	119-E-A-R-Q.....	D-S.....	..D G V D H R T V S E I I K L K D P D T.....	145
<i>Noccaea caenulensis</i> [L_T17688_c0_g1_l1_g_2717/1-213	119-E-A-R-Q.....	D-S.....	..D G V D H R T V S E I I K L K D P D T Q.....	146
<i>Noccaea caenulensis</i> [GA_T11053_c0_g1_l1_g_34165/1-213	119-E-A-R-Q.....	D-S.....	..D G V D H R T V S E I I K L K D P D T.....	145
<i>Noccaea caenulensis</i> [L_T17411_c0_g1_l1_g_56299/1-213	119-E-A-R-Q.....	D-S.....	..D G V D H R T V S E I I K L K D P D T.....	145
<i>Rosa chinensis</i> [RchOBHb_Chr3g046096/1-229	129-Q-L--R-E.....	D-S.....	..D O L D H R A V S E V L A K L K D P D T.....	154
<i>Prunus pennsylvanica</i> PRUPE_6G290000/1-253	140-Q-L--R-E.....	D-S.....	..D O L D H R A V S E V L V K L K D P D T.....	165
<i>Prunus dulcis</i> [ALMOND_28028996/1-240	140-Q-L--R-E.....	D-S.....	..D O L D H R A V S E V L V K L K D P D T.....	165
<i>Malus domestica</i> [DDB4_036312/1-296	201-Q-F--K-E.....	D-S.....	..D O L D H R A V S E V L V K L K D P D T.....	226
<i>Malus baccata</i> [C1H4H_040009/1-241	142-Q-F--K-E.....	D-G.....	..D O L D H R A V S E V L V K L K D P D T.....	167
<i>Trema orientalis</i> [TorR033c02_098860/1-233	137-Q-L--R-E.....	D-S.....	..D O L D H H A V E E V L T K L K D P D T.....	162
<i>Parapania andersoni</i> [PanWU01c14_361630/1-240	144-Q-V--H-E.....	D-S.....	..D O L D H H A V E E V L T K L K D P D T.....	169
<i>Rhizophora mucronata</i> [WJQ400303244/1-238	148-S-E-I-Q-E.....	D-S.....	..D O L D H H A V S E V L I K L K D P D T.....	174
<i>Populus alba</i> [P0086_0000505270/1-240	146-K-H--Q-E.....	D-S.....	..D S I D Q H A I S E V L V K L K D P D T.....	169
<i>Populus trichocarpa</i> [POPTT_0160133400/1-242	146-K-H--Q-E.....	D-S.....	..D S I D Q H A I S E V L V K L K D P D T.....	171
<i>Populus trichocarpa</i> [POPTT_006G107300/1-242	146-K-H--Q-E.....	D-S.....	..D S I D Q H A I S E V L V K L K D P D T.....	171
<i>Juglans regia</i> [LOC1089998/1-249	144-Q-L--R-E.....	D-S.....	..D O L D H R T V S E I I V K L K D P D T.....	174
<i>Juglans regia</i> [LOC10901925/1-244	140-Q-F--H-E.....	D-S.....	..D O L D H R A V S E I L V K L K D P D T.....	169
<i>Fagus sylvatica</i> [F38_LOCUS40270/1-209	115-Q-L--H-E.....	D-S.....	..D O L D H R V F S E V L A K L K D P D T.....	140
<i>Cucumis sativus</i> [C030396/1-233	146L-F			

192

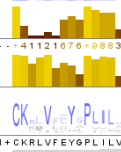
193

Conservation 
Quality 
Consensus 
Occupancy 

194

Oryza_nivara|NA|A0A0E0FGR9|1-226 178 CKMMLENTPLILV 191
Oryza_sativa_subap_indica|Cel_00546|1-262 214 CKMMLENTPLILV 227
Oryza_sativa_subap_indica|Cel_34843|1-245 194 CKRMMLQVPLVLV 207
Oryza_sativa_subap_indica|Cel_37293|1-245 194 CKRMMLQVPLVLV 207
Oryza_sativa_subap_indica|Cel_19500|1-223 174 CKKEIQNAPFILE 187
Oryza_sativa_subap_japonica|Cel_3p112|200|1-245 194 CKRMMLQVPLVLV 207
Oryza_sativa_subap_japonica|P00286|10_2|1-240 192 CKMMLENTPLILV 205
Oryza_sativa_subap_japonica|Cel_0p166700|1-226 178 CKMMLENTPLILV 191
Oryza_sativa_subap_japonica|Cel_05g0334400|1-223 174 CKKEIQNAPFILE 187
Brachypodium_distachyon|BRADL_4g44500v3|1-246 197 CKRMMLQVPLILV 210
Brachypodium_distachyon|BRADL_4g5560v3|1-245 194 CKRMMLQVPLILV 207
Brachypodium_distachyon|BRADL_3p1070v3|1-236 187 CKMLIQNAPVILV 200
Brachypodium_distachyon|BRADL_3p04110v3|1-235 187 CKRLVLENTPLILV 200
Hordeum_vulgare_subap_vulgare|NA|A0A267W79|1-318 267 CKRLVLENTPLILV 280
Hordeum_vulgare_subap_vulgare|NA|A0A267R2L5|1-266 215 CKRMMLQVPLILV 228
Hordeum_vulgare_subap_vulgare|NA|F2DBE9|1-246 195 CKRMMLQVPLILV 208
Hordeum_vulgare_subap_vulgare|NA|F2C0A89|1-241 190 CKRLVLENTPLILV 203
Hordeum_vulgare_subap_vulgare|NA|A0A267KA24|1-238 190 CKRLVLENTPLILV 203
Aegilops_tauschii_subap_stragulata|NA|A0A453HW0|1-260 200 CKRLVLENTPLILV 222
Aegilops_tauschii|F775_31720|1-255 204 CKRLVLENTPLILV 217
Aegilops_tauschii_subap_stragulata|NA|A0A453E3V5|1-235 190 CKMMLENTPLILV 203
Aegilops_tauschii_subap_stragulata|NA|A0A4522QF0|1-224 175 CKMMIQNAPFILE 188
Triticum_aestivum|NA|A0A368JCF3|1-255 204 CKRLVLENTPLILV 217
Triticum_aestivum|NA|A0A368SKEQ|2|1-249 198 CKRMMLQVPLILV 211
Triticum_aestivum|NA|A0A368LJX9|1-249 198 CKMMLENTPLILV 211
Triticum_aestivum|NA|A0A368MMT5|1-246 195 CKRMMLQVPLILV 208
Triticum_aestivum|NA|A0A368BRE3|1-240 187 CKRLVLENTPLILV 202
Triticum_aestivum|NA|A0A368E64|1-238 193 CKRLVLENTPLILV 206
Triticum_aestivum|NA|A0A368SV5L6|1-235 189 CKMMIQNAPVILE 203
Triticum_aestivum|NA|A0A368FHS9|1-233 188 CKRLVLENTPLILV 201
Triticum_aestivum|NA|A0A368A1D|1|1-209 190 CKMMIQNAPFILE 173
Triticum_aestivum|NA|A0A368Z46|1|1-227 178 CKMMIQNAPFILE 191
Triticum_turgidum_subap_dunum|TRTD_1AvTG20820|1-283 175 CKMMIQNAPFILE 188
Triticum_turgidum_subap_dunum|TRTD_4Bv1G046790|1-256 205 CKRLVLENTPLILV 219
Triticum_turgidum_subap_dunum|TRTD_4AvTG152380|1-255 204 CKRLVLENTPLILV 217
Triticum_turgidum_subap_dunum|TRTD_3AvTG29160|1-253 208 CKRLVLENTPLILV 221
Triticum_turgidum_subap_dunum|TRTD_5AvTG112490|1-249 198 CKRMMLQVPLILV 211
Triticum_turgidum_subap_dunum|TRTD_5BvTG093140|1-249 198 CKRLVLENTPLILV 211
Triticum_turgidum_subap_dunum|TRTD_1BvTG194670|1-244 190 CKMMIQNAPFILE 191
Triticum_turgidum_subap_dunum|TRTD_3BvTG033240|1-235 190 CKRLVLENTPLILV 203
Triticum_uranu|TRUR3_03527|1-283 175 CKMMIQNAPFILE 188
Triticum_uranu|TRUR3_22517|1-255 204 CKRLVLENTPLILV 217
Triticum_uranu|TRUR3_29270|1-253 208 CKRLVLENTPLILV 221
Triticum_uranu|TRUR3_22718|1-246 195 CKRMMLQVPLILV 208
Anundo_donas|NA|A0A048979|1-250 193 CKKEIQNAPFILE 207
Anundo_donas|NA|A0A0489V1|2|1-229 181 CKRLVLENTPLILV 194
Anundo_donas|NA|A0A0489V0R9|1-222 173 CKKEIQNAPFILE 186
Anundo_donas|NA|A0A0489V254|1-191 135 CKKEIQNAPFILE 148
Anundo_donas|NA|A0A0489QW3|1-166 110 CKKEIQNAPFILE 123
Anundo_donas|NA|A0A0489QWV1|1-122 74 CKRLVLENTPLILV 87
Eragrostis_cervina|EJB05_31312|1-320 199 CKRMMLQVPLILV 206
Eragrostis_cervina|EJB05_03028|1-246 195 CKRLVLENTPLILV 208
Eragrostis_cervina|EJB05_03037|1-240 189 CKRLVLENTPLILV 202
Eragrostis_cervina|EJB05_29979|1-225 195 CKRLVLENTPLILV 208
Eragrostis_cervina|EJB05_34950|1-223 174 CKKEIQNAPFILE 187
Eragrostis_cervina|EJB05_29935|1-176 124 CKRLVLENTPLILV 137
Eragrostis_cervina|EJB05_31440|1-107 56 CKRLVLENTPLILV 69
Sorghum_bicolor|SORB1_3008G032600|1-247 196 CKRLVLENTPLILV 206
Sorghum_bicolor|SORB1_3003G055700|1-227 179 CKRLVLENTPLILV 192
Sorghum_bicolor|SORB1_3009G097200|1-180 158 CKRLVLENTPLILV 165
Zea_mays|Zm0014fa_045690|1-237 185 VGI V D I F A I L H G F P D N V R T N Q K G F V W A I H R S M P Y R L M S C H M Z 228
Zea_mays|ZEAAMB73_2m0000102337|1|1-239 177 CKRLVLENTPLILV 220
Zea_mays|ZEAAMB73_2m00001042734|1-240 185 CKRLVLENTPLILV 198
Zea_mays|ZEAAMB73_2m00001039719|1-229 181 CKRLVLENTPLILV 194
Dichanthelium_oligoanthum|BAE44_0015218|1-291 200 VRNSVLI SYDPSMCKLFLKIAEITRSCFLLLHLMYMNCCO CKRLVLENTPLILV 263
Dichanthelium_oligoanthum|BAE44_0009503|1-182 136 CKRLVLENTPLILV 145
Paniceum_hallii_var_hallii|GQ55_BG009100|1-237 186 CKRLVLENTPLILV 199
Paniceum_hallii_var_hallii|GQ55_BG490800|1-232 194 CKRLVLENTPLILV 197
Paniceum_hallii_var_hallii|GQ55_3G0007700|1-229 179 CKRLVLENTPLILV 192
Paniceum_hallii_var_hallii|GQ55_3G333500|1-224 175 CKKEIQNAPFILE 188
Paniceum_miliaceum|C2845_PM0604340|1-263 215 CKRLVLENTPLILV 228
Paniceum_miliaceum|C2845_PM17000420|1-231 180 CKRLVLENTPLILV 193
Paniceum_miliaceum|C2845_PM05021750|1-204 155 CKRLVLENTPLILV 168
Paniceum_miliaceum|C2845_PM07000670|1-203 135 CKRLVLENTPLILV 168
Paniceum_miliaceum|C2845_PM06026640|1-184 155 CKRLVLENTPLILV 148
Setaria_italica|SETIT_7G337400v2|1-230 179 CKRLVLENTPLILV 192
Setaria_italica|SETIT_5G117400v2|1-227 175 CKRLVLENTPLILV 188
Setaria_italica|SETIT_BG015000v2|1-226 175 CKRLVLENTPLILV 188
Setaria_italica|SETIT_3G284400v2|1-225 176 CKRLVLENTPLILV 188
Setaria_vindis|SEVIR_7G337200v2|1-230 179 CKRLVLENTPLILV 192
Setaria_vindis|SEVIR_5G113900v2|1-227 179 CKRLVLENTPLILV 192
Setaria_vindis|SEVIR_BG013800v2|1-226 175 CKRLVLENTPLILV 188
Setaria_vindis|SEVIR_3G292600v2|1-225 176 CKRLVLENTPLILV 188
Aquillegia_cornuta|AQUCO_00400489v1|1-223 170 CKRLVLENTPLILV 183
Macleaya_cordata|BVCB0_1837g47|1-262 232 CEELLKKYFPLILV 245
Macleaya_cordata|BVCB0_9017g10|1-212 197 CKRLVLENTPLILV 172
Papaver_nominatum|C6167_002404|1-357 209 CKRLVLENTPLILV 207
Nelumbo_nucifera|LOC10460198|1-243 193 CKRLVLENTPLILV 206
Nelumbo_nucifera|LOC104602464|1-238 189 CKRLVLENTPLILV 202
Spinacia_oleracea|SOVIF_050110|1-231 178 CKRMMLQVPLILV 191
Actinidia_chinensis_var_chinensis|CEY00_Acc01859|1-258 204 CKRLVLENTPLILV 217
Actinidia_chinensis_var_chinensis|CEY00_Acc00070|1-246 194 CKRLVLENTPLILV 207
Actinidia_chinensis_var_chinensis|CEY00_Acc08729|1-244 192 CKRLVLENTPLILV 205
Actinidia_chinensis_var_chinensis|CEY00_Acc01859|1-227 173 CKRLVLENTPLILV 186
Davidia_involucrata|Din_006700|1-269 187 CKRLVLENTPLILV 200
Davidia_involucrata|Din_026378|1-245 193 CKRLVLENTPLILV 206
Nyssa_chinensis|F0562_017856|1-239 103 CKRLVLENTPLILV 206
Nyssa_chinensis|F0562_018153|1-125 94 CKRLVLENTPLILV 107
Asteriscus_annuus|CT12_AA263659|1-132 189 CKRLVLENTPLILV 166
Asteriscus_annuus|CT12_AA249499|1-198 153 CKRLVLENTPLILV 166
Helianthus_annuus|HannXRX_Chrl0g028629|1-181 135 CKRLVLENTPLILV 148
Cymara_cardunculus_var_acylasmus|Ccd_003008|1-231 185 CKRLVLENTPLILV 166
Lactuca_sativa|LAT_9X3806|1|1-229 186 CKRLVLENTPLILV 199
Daucus_carota_subap_sativus|DCAR_010960|1-241 189 CKRMMLQVPLILV 202
Daucus_carota_subap_sativus|DCAR_010965|1-143 104 CKRLVLENTPLILV 117
Daucus_carota_subap_sativus|DCAR_010965|1-143 104 CKRLVLENTPLILV 117
Daucus_carota_subap_sativus|DCAR_010965|1-143 104 CKRLVLENTPLILV 117
Erythranthe_guttata|MMGLU_ngvfa013247ng|1-226 180 CKRLVLENTPLILV 162
Genlisea_aurea|JMS68_00799|1-243 183 CKRLVLENTPLILV 166
Hemianthus_angelicoides|COL12_116059|1-229 177 CKRLVLENTPLILV 160
Stipa_sibirica|STAS_21183|1-236 183 CKRLVLENTPLILV 166
Stipa_sibirica|STAS_23432|1-199 160 CKRLVLENTPLILV 166
Coffea_cannabina|GSCOG_70002323400|1-294 241 CKRMMLQVPLILV 273
Cuscuta_australis|DM860_002763|1-226 180 CKRLVLENTPLILV 193
Cuscuta_campetris|CCAM_LOCUS310859|1-226 180 CKRLVLENTPLILV 193
Cuscuta_campetris|CCAM_LOCUS31789|1-223 177 CKRLVLENTPLILV 190
Nicotiana_glabra|NA4A9_38789|1-245 193 CKRLVLENTPLILV 205
Nicotiana_glabra|NA4A9_19559|1-238 193 CKRLVLENTPLILV 190
Nicotiana_glabra|LOC104216406|1-245 192 CKRLVLENTPLILV 205
Nicotiana_glabra|LOC104224609|1-238 186 CKRLVLENTPLILV 190
Nicotiana_glabra|LOC107812754|1-245 192 CKRLVLENTPLILV 205
Nicotiana_glabra|LOC10781660|1-245 192 CKRLVLENTPLILV 205
Nicotiana_glabra|LOC10781660|1-238 186 CKRLVLENTPLILV 190
Nicotiana_glabra|LOC10777346|1-238 186 CKRLVLENTPLILV 190
Capsicum_annuum|LOC107784342|1-241 189 CKRMMLQVPLILV 202
Capsicum_annuum|LOC107851224|1-124 189 CKRMMLQVPLILV 202
Capsicum_baccatum|CQW23_24170|1-241 189 CKRMMLQVPLILV 202
Capsicum_baccatum|CQW23_22278|1-197 189 CKRMMLQVPLILV 202
Capsicum_baccatum|CQW23_29495|1-163 189 CKRMMLQVPLILV 202
Capsicum_chinense|BC332_26027|1-241 189 CKRMMLQVPLILV 202
Capsicum_chinense|BC332_31415|1-121 189 CKRMMLQVPLILV 202
Solanum_chacoense|NA|A0A0V0H4M7|1-278 190 CKRLVLENTPLILV 203
Solanum_chacoense|NA|A0A0V0H4M7|1-278 190 CKRLVLENTPLILV 202
Solanum_inubacuum|LOC10603503|1-242 190 CKRLVLENTPLILV 203
Solanum_inubacuum|NA|A0A3Q0700|1-238 187 CKRLVLENTPLILV 200
Vitis_vinifera|VT_08e0058y01030|1-278 237 CKRMMLQVPLILV 250
Vitis_vinifera|VITSV_040420|1-239 198 CKRMMLQVPLILV 211
Vitis_vinifera|Paapl_1|1-195 154 CKRMMLQVPLILV 167
Vitis_vinifera|VITSV_040420|1-174 125 CKRMMLQVPLILV 136
Vitis_vinifera|CK203_030312|1-150 125 CKRMMLQVPLILV 136

Occupanc



197

Oryza_piava	[NA]A0A06GFR91-226	219TMLLSAAS.....	226
Oryza_sati	subap_indica[Oel_00546/1-262	255TMLLSAAS.....	262
Oryza_sati	subap_indica[Oel_34843/1-245	241LV-RDA.....	245
Oryza_sati	subap_indica[Oel_37293/1-245	241LV-RDA.....	245
Oryza_sati	subap_indica[Oel_19509/1-223	221LV-RDA.....	223
Oryza_sati	subap_japonica[Oel_0112200/1-245	241LV-RDA.....	245
Oryza_sati	subap_japonica[P00286/10-2/1-240	233TMLLSAAS.....	240
Oryza_sati	subap_japonica[Oel_010166700/1-226	219TMLLSAAS.....	226
Oryza_sati	subap_japonica[Oel_050334400/1-223	221SAA.....	223
Brachypodium_distachyon	[BRADL_4g44500v3/1-248	244LRS EA.....	248
Brachypodium_distachyon	[BRADL_4g45500v3/1-245	241LRS EA.....	245
Brachypodium_distachyon	[BRADL_3g31070v3/1-236	234LRS EA.....	236
Brachypodium_distachyon	[BRADL_3g04110v3/1-235	228TVLSAAL.....	235
Hordeum_vulgare	subap_vulgare[NA]A0A267M79/1-318	314LLRDE.....	318
Hordeum_vulgare	subap_vulgare[NA]A0A267R2L5/1-266	262LRS DA.....	266
Hordeum_vulgare	subap_vulgare[NA]A20B69/1-246	242LRS DA.....	246
Hordeum_vulgare	subap_vulgare[NA]A20C49/1-241	237LLRDE.....	241
Hordeum_vulgare	subap_vulgare[NA]A0A267K424/1-238	231TIPLSATL.....	238
Aegilops_tauschii	subap_strangulata[NA]A0A453HW0/1-260	256LLRDA.....	260
Aegilops_tauschii	[F775_31720/1-255	251LLRDA.....	255
Aegilops_tauschii	subap_strangulata[NA]A0A453E3V5/1-235	231LSAAL.....	235
Aegilops_tauschii	subap_strangulata[NA]A0A452QF0/1-224	222STA.....	224
Triticum_aestivum	[NA]A0A384JC2/1-255	251LLRDA.....	255
Triticum_aestivum	[NA]A0A384KEQ2/1-249	245LRS DA.....	249
Triticum_aestivum	[NA]A0A384LJX3/1-249	245LRS DA.....	249
Triticum_aestivum	[NA]A0A384MMT5/1-246	242LRS DA.....	246
Triticum_aestivum	[NA]A0A384BRE3/1-240	236LLRDA.....	240
Triticum_aestivum	[NA]A0A384E4/1-238	234LSAAL.....	238
Triticum_aestivum	[NA]A0A384V5L6/1-235	233YTA.....	235
Triticum_aestivum	[NA]A0A384FHS9/1-233	229LSAAL.....	233
Triticum_aestivum	[NA]A0A384A1D/1-209	207STA.....	209
Triticum_aestivum	[NA]A0A384S246/1-227	225STV.....	227
Triticum_turgidum	subap_dunali[TRTD_14v1G205500/1-293	291	WPCIDYAFVTEIFATLIISEEPWYFARTMFDALKGCTMMDLSVGWLNWYVLF6	293
Triticum_turgidum	subap_dunali[TRTD_4Bv1G049790/1-256	252LLRDA.....	256
Triticum_turgidum	subap_dunali[TRTD_4Av1G152380/1-255	251LLRDA.....	255
Triticum_turgidum	subap_dunali[TRTD_3Av1G029160/1-253	249LSAAL.....	253
Triticum_turgidum	subap_dunali[TRTD_5Av1G112490/1-249	245LRS DA.....	249
Triticum_turgidum	subap_dunali[TRTD_5Bv1G093140/1-249	246LRS DA.....	249
Triticum_turgidum	subap_dunali[TRTD_1Bv1G194670/1-244	239LWRSVYIITVVK.....	244
Triticum_turgidum	subap_dunali[TRTD_3Bv1G033240/1-235	231LSAAL.....	235
Triticum_urartu	[TRUR3_03257/1-283	231	WPCIDYAFVTEIFATLIISEEPWYFARTMFDALKGCTMMDLSVGWLNWYVLF6	283
Triticum_urartu	[TRUR3_22517/1-255	251LLRDA.....	255
Triticum_urartu	[TRUR3_29270/1-253	249LSAAL.....	253
Triticum_urartu	[TRUR3_22719/1-246	242LRS DA.....	246
Anundo_donas	[NA]A0A049R79/1-250	241LLPELPRTRA.....	250
Anundo_donas	[NA]A0A049RVI2/1-229	222SMPLSATL.....	229
Anundo_donas	[NA]A0A049V0R9/1-222	220SAA.....	222
Anundo_donas	[NA]A0A049V254/1-191	182LLPELPRTRA.....	191
Anundo_donas	[NA]A0A049QW3/1-166	157LLPELPRTRA.....	166
Anundo_donas	[NA]A0A049Q8/1-122	115SMPLSATL.....	122
Eragrostis_curuia	[EJB05_31312/1-320	314SVLSAS.....	320
Eragrostis_curuia	[EJB05_03028/1-246	242LLNDA.....	246
Eragrostis_curuia	[EJB05_03037/1-240	236LLS DA.....	240
Eragrostis_curuia	[EJB05_29979/1-225	221STA.....	225
Eragrostis_curuia	[EJB05_34950/1-223	221STA.....	223
Eragrostis_curuia	[EJB05_29835/1-176	171LLRDA.....	176
Eragrostis_curuia	[EJB05_31440/1-107	103LLS DA.....	107
Sorghum_bicolor	[SORB1_3008G032600/1-247	243LLS DA.....	247
Sorghum_bicolor	[SORB1_3003G055700/1-227	220TMPLSATL.....	227
Sorghum_bicolor	[SORB1_3009G097200/1-180	171TMPLSATL.....	180
Zea_mays	[Zm0014a_038550/1-277	277SVLKAALLLS DA.....	277
Zea_mays	[Zm0014a_04459/1-234	222SVLKAALLLS DA.....	234
Zea_mays	[ZmAMMB73_2m0000102337/1-239	228SVLKAALLLS DA.....	239
Zea_mays	[ZmAMMB73_2m00001042734/1-240	232LWSDAWLEO.....	240
Zea_mays	[ZmAMMB73_2m00001039719/1-229	222AMPLSAML.....	229
Dichanthelium_oligosanthos	[BAE4_0015216/1-291	287LVNDA.....	291
Dichanthelium_oligosanthos	[BAE4_0008106/1-227	220SMPLSATL.....	227
Dichanthelium_oligosanthos	[BAE4_0009052/1-182	176VMKRFV.....	182
Panicum_halli	var_halli[GQ55_BG009100/1-237	233LVSDA.....	237
Panicum_halli	var_halli[GQ55_BG490800/1-232	225SMPLSATL.....	232
Panicum_halli	var_halli[GQ55_3G0007700/1-229	225LLS DA.....	229
Panicum_halli	var_halli[GQ55_3G333500/1-224	222SSV.....	224
Panicum_miliaceum	[C2845_PM0050430/1-263	256SMPLSATL.....	263
Panicum_miliaceum	[C2845_PM17000420/1-231	227LVSDA.....	231
Panicum_miliaceum	[C2845_PM05021750/1-204	202SSA.....	204
Panicum_miliaceum	[C2845_PM07G00670/1-203	196GMPLSATL.....	203
Panicum_miliaceum	[C2845_PM06G26640/1-184	182SSV.....	184
Setaria_italica	[SETT_7G337400/1-230	226LLS DA.....	230
Setaria_italica	[SETT_8G11400/1-227	222NIPLSATL.....	227
Setaria_italica	[SETT_BG015000/1-226	222LVSDA.....	226
Setaria_italica	[SETT_3G284400/1-225	223SSA.....	225
Setaria_vindis	[SEVR_7G337200/1-230	226LLS DA.....	230
Setaria_vindis	[SEVR_5G113900/1-227	220SIPLSATL.....	227
Setaria_vindis	[SEVR_BG138000/1-226	222LVSDA.....	226
Setaria_vindis	[SEVR_3G292600/1-225	223SSA.....	225
Aquilegia_scoenicea	[AQUCO_00400489v/1-223	217KSLVADA.....	223
Macleaya_cordata	[BVCB0_1837g47/1-262	262TSLLAES.....	262
Macleaya_cordata	[BVCB0_9017g10/1-212	204TSLLAES.....	212
Papaver_somniferum	[C6167_002404/1-357	346SFSTSAETSLLAAS.....	357
Nelumbo_nucifera	[LOC104597198/1-243	236FYGGTILII.....	243
Nelumbo_nucifera	[LOC104602464/1-228	220SV-EGRIEMV-TSS.....	228
Spinacia_oleracea	[SOVF_050110/1-231	220SV-EGRIEMV-TSS.....	231
Actinidia_chinensis	var_chinensis[CEV00_Acc01859/1-258	246PVVAKPPMAMII-SV.....	258
Actinidia_chinensis	var_chinensis[CEV00_Acc0072/1-246	239ETLISAM.....	246
Actinidia_chinensis	var_chinensis[CEV00_Acc08725/1-244	238TLTILII.....	244
Actinidia_chinensis	var_chinensis[CEV00_Acc01859/1-227	220PPMAMII-SV.....	227
Davidia_involucrata	[Dm_006700/1-269	247	PMIYGRFFTAIVICLLLVSSHHPW	269
Davidia_involucrata	[Dm_026378/1-245	240SLF-SES.....	245
Nyssa_chinensis	[F0562_017859/1-239	239SLF-SES.....	239
Nyssa_chinensis	[F0562_018153/1-225	225SLF-SES.....	225
Atenuia_annua	[CT12_AA38550/1-132	132TLLSDH.....	132
Atenuia_annua	[CT12_AA39490/1-199	193DTLLSDH.....	199
Helianthus_annuus	[HannXRO_Chrl0g028629/1-181	174EASLVSDN.....	181
Cynara_scolymus	var_scolymus[CCid_003008/1-231	225ASKISDN.....	231
Lachua_sati	[SAT_9X3806/1-229	225LISDN.....	229
Desouz_cariota	subap_sathus[CAPL_010900/1-241	226QASTOETIML-AS.....	241
Desouz_cariota	subap_sathus[CAPL_010655/1-143	143QASTOETIML-AS.....	143
Donocerae_hygoneticum	[F511_29468/1-190	178ALIRTDSTLH-SAS.....	190
Erythranthe_guttata	[MMGU_ngrv1a013247ng/1-226	226YIRSNHAQSG.....	226
Genilaea_aurea	[M569_00799/1-243	234YIRSNHAQSG.....	243
Hemidranthus_papilionaceus	[COL12_116059/1-229	218VAGTOMPTHAAS.....	229
Shipa_asiatica	[STAS_21183/1-236	226DSFMHAAS.....	236
Shipa_asiatica	[STAS_33432/1-199	199DSFMHAAS.....	199
Coffea_caneophora	[GSCOG_70002323400/1-294	287ETSLRAAS.....	294
Cuscuta_australis	[DM860_002763/1-226	220ASMHFAS.....	226
Cuscuta_campestris	[CCAM_LOCUS31085/1-226	226ASMHFAS.....	226
Cuscuta_campestris	[CCAM_LOCUS32789/1-223	223ASMHFAS.....	223
Nicotiana_attenuata	[A4A9_28789/1-245	239ASPKMOTSMHSAS.....	245
Nicotiana_attenuata	[A4A9_19559/1-238	226DLPKMOTSMHSAS.....	238
Nicotiana_sylvestris	[LOC104216406/1-245	231LQASPKMOTSMHSAS.....	245
Nicotiana_sylvestris	[LOC104224609/1-238	226VLPKMOTSMHSTS.....	238
Nicotiana_tabacum	[LOC107812754/1-245	231LQASPKMOTSMHSAS.....	245
Nicotiana_tabacum	[LOC10780970/1-245	231LQASPKMOTSMHSAS.....	245
Nicotiana_tabacum	[LOC107792809/1-238	226VLPKMOTSMHSAS.....	238
Nicotiana_tabacum	[LOC10777346/1-238	226VLPKMOTSMHSAS.....	238
Capiscum_annuum	[LOC107843427/1-241	235SSLSAS.....	241
Capiscum_annuum	[LOC107851224/1-124	124SSLSAS.....	124
Capiscum_baccatum	[CQW23_24170/1-241	235SSLSAS.....	241
Capiscum_baccatum	[CQW23_22279/1-197	197SSLSAS.....	197
Capiscum_baccatum	[CQW23_29496/1-163	163SSLSAS.....	163
Capiscum_chinense	[BC332_26027/1-241	235SSLSAS.....	241
Capiscum_chinense	[BC332_31415/1-121	121SSLSAS.....	121
Solanum_chacoense	[NA]A0A0V0HM7/1-278	278TSRPRKQTSLSH-SASXGKNNGIQTRLGTGYCCVSKQLNIIKVVYEILCIMYII	278
Solanum_chacoense	[NA]A0A0V0HW/1-273	273ALPAXQTSLSH-SASXGKNNGIQTRLGTGYCCVSKQLNIIKVVYEILCIMYII	273
Solanum_habrosum	[S0360353/1-241	227SLH-SAS.....	241
Solanum_ipoericum	[NA]A0A3C0700/1-238	227AS-RKQTSLSH-SAS.....	238
Vitis_vinifera	[WT_08e0058y01030/1-278	278AS-RKQTSLSH-SAS.....	278
Vitis_vinifera	[WTSV_040420/1-239	239AS-RKQTSLSH-SAS.....	239
Vitis_vinifera	[Paapl_1_1-195	195AS-RKQTSLSH-SAS.....	195
Vitis_vinifera	[WTSV_040420/1-174	166EAAAS-SS.....	174
Vitis_vinifera	[CK203_030312/1-150	150EAAAS-SS.....	150

<i>Vitis_riparia</i> [NA Q9M614 1-174	172	-----SDS-----	174
<i>Arachis_hypogaea</i> [Ahy_B04g071408 1-245	243	-----GAY-----	245
<i>Arachis_hypogaea</i> [Ahy_A03g006469 1-217	210	-----ETFLFSDS-----	217
<i>Arachis_hypogaea</i> [Ahy_B03g001901 1-217	211	-----TLFLFSDS-----	217
<i>Arachis_hypogaea</i> [Ahy_A04g019829 1-212	207	-----ALL-SDS-----	212
<i>Lupinus_angustifolius</i> [Tanj G_22489 1-222	218	-----AFL-SF-----	222
<i>Lupinus_angustifolius</i> [Tanj G_19379 1-204	197	-----ELPLVSDS-----	204
<i>Cicer_arietinum</i> [LOC101491522 1-279	269	-----RMLHLKIKALY-----	279
<i>Cicer_arietinum</i> [LOC101508250 1-215	209	-----IFPML-SDS-----	215
<i>Medicago_truncatula</i> [MTR_7g072550 1-242	223	-----IFVGNIFRMVQLLK-IF-----	242
<i>Medicago_truncatula</i> [MtrunA17_Chrg001314 1-223	217	-----TALFSDS-----	223
<i>Medicago_truncatula</i> [MTR_029040 1-215	209	-----TALF-SDS-----	215
<i>Trifolium pratense</i> [L195_g026334 1-194	188	-----IPML-SDS-----	194
<i>Trifolium_subterraneum</i> [TSLD_266660 1-215	210	-----TML-SDS-----	215
<i>Trifolium_subterraneum</i> [TSLD_160390 1-181	179	-----LAW-----	181
<i>Lotus_japonicus</i> [WJ3399 1-216	210	-----IFPLL-SDS-----	216
<i>Cajanus_cajan</i> [KK1_035920 1-221	216	-----AFL-SDS-----	221
<i>Cajanus_cajan</i> [KK1_029931 1-217	211	-----VPLI-SDS-----	217
<i>Mucuna pruriens</i> [CR513_15124 1-288	283	-----CRFL-SDS-----	288
<i>Mucuna pruriens</i> [CR513_55238 1-242	236	-----VPLI-SDS-----	242
<i>Phaseolus_vulgaris</i> [PHAVL_0865094900 1-222	217	-----AFL-SDS-----	222
<i>Phaseolus_vulgaris</i> [PHAVL_0865094700 1-217	213	-----LLWKQ-----	217
<i>Glycine_max</i> [GLYMA_18G121100 1-250	231	RSNRFVVYLLQKIAIYITTH-----	250
<i>Glycine_max</i> [GLYMA_19G111400 1-237	218	RSNRFVVYLLQKIAIYITTO-----	237
<i>Glycine_max</i> [GLYMA_09G277100 1-237	218	RSNRFVYLLQKIAIYITTO-----	237
<i>Glycine_max</i> [GLYMA_18G211900 1-236	231	-----AFLSVS-----	236
<i>Glycine_max</i> [GLYMA_01G113100 1-216	204	-----EASIM-EVPLI-SDS-----	216
<i>Glycine_max</i> [GLYMA_09G277200 1-212			
<i>Glycine_max</i> [GLYMA_04G159500 1-202	183	HSNIFVVYLLQKISIIYITTH-----	202
<i>Glycine_max</i> [GLYMA_19G111500 1-181			
<i>Glycine_ega</i> [DOYG5_025469 1-265			
<i>Glycine_ega</i> [DOYG5_001390 1-253			
<i>Glycine_ega</i> [DOYG5_025469 1-237	241	-----EASIM-EVPLI-SDS-----	253
<i>Glycine_ega</i> [DOYG5_049180 1-229	218	RSNRFVYLLQKIAIYITTO-----	229
<i>Vigna_angularis_var_angularis</i> [Vigan_04G117700 1-252	224	-----AFLSVS-----	252
<i>Vigna_angularis_var_angularis</i> [Vigan_04G117800 1-219	230	-----VSSLQKIAKYI-TPY-----	219
<i>Vigna_angularis_var_angularis</i> [Vigan_09G109000 1-217	214	-----AFL-SDS-----	217
<i>Vigna_radiata_var_radiata</i> [LOC10678179 1-219	204	-----EASIMEEPLI-SDS-----	219
<i>Vigna_radiata_var_radiata</i> [LOC106784929 1-217	210	-----EPLIFDS-----	217
<i>Vigna_radiata_var_radiata</i> [LOC106768949 1-211	208	-----AFL-SDS-----	211
<i>Vigna_unguiculata</i> [DE072_LG10g3244 1-238	225	-----VYSLQKIAKYI-ATH-----	238
<i>Vigna_unguiculata</i> [DE072_LG10g3245 1-220	215	-----AFL-SDS-----	220
<i>Vigna_unguiculata</i> [DE072_LG8g11591 1-217	204	-----EASIMEEPLI-SDS-----	217
<i>Citrus_unkou</i> [CUMW_001149 1-204			
<i>Acer_yangbiense</i> [EZV62_016774 1-186			
<i>Eucalyptus_grandis</i> [EUGRSUZ_K01273 1-227	216	-----VPT-KE-MPLL-SDS-----	227
<i>Eucalyptus_grandis</i> [EUGRSUZ_A00687 1-219	215	-----VVASS-----	219
<i>Punica_granatum</i> [CRG98_041613 1-272	260	-----LPT-SQEAUVL-SDS-----	272
<i>Punica_granatum</i> [CRG98_041612 1-227	215	-----LPT-TEEVVL-SYR-----	227
<i>Punica_granatum</i> [CRG98_016880 1-220	220	-----VOLLST-----	220
<i>Conthosus_capulata</i> [CGCVL1_28877 1-226	223	-----L-ADS-----	226
<i>Conthosus_olitorius</i> [COL04_30004 1-226	223	-----L-ADS-----	226
<i>Gossypium_ardoreum</i> [F383_27015 1-233	230	-----V-ADS-----	233
<i>Gossypium_ardoreum</i> [F383_21360 1-227	224	-----V-ADS-----	227
<i>Gossypium_barbadense</i> [GGBAR_AA12144 1-247	245	-----V-ADS-----	247
<i>Gossypium_barbadense</i> [GGBAR_AA02853 1-247	244	-----V-ADS-----	247
<i>Gossypium_barbadense</i> [ES319_D10G128500v 1-233	230	-----V-ADS-----	233
<i>Gossypium_barbadense</i> [ES319_A10G160500v 1-233	230	-----V-ADS-----	233
<i>Gossypium_barbadense</i> [ES319_D02G005800v 1-227	224	-----V-ADS-----	227
<i>Gossypium_barbadense</i> [ES319_A02G004900v 1-227	224	-----V-ADS-----	227
<i>Gossypium_darwinii</i> [ES288_D10G138900v 1-233	230	-----V-ADS-----	233
<i>Gossypium_darwinii</i> [ES288_A10G179500v 1-233	230	-----V-ADS-----	233
<i>Gossypium_darwinii</i> [ES288_A02G005100v 1-227	224	-----V-ADS-----	227
<i>Gossypium_darwinii</i> [ES288_D02G001700v 1-227	224	-----V-ADS-----	227
<i>Gossypium_hirsutum</i> [LOC107896756 1-233	230	-----V-ADS-----	233
<i>Gossypium_hirsutum</i> [LOC107819549 1-233	230	-----V-ADS-----	233
<i>Gossypium_hirsutum</i> [LOC10783569 1-227	224	-----V-ADS-----	227
<i>Gossypium_hirsutum</i> [LOC107903579 1-227	224	-----V-ADS-----	227
<i>Gossypium_mustelinum</i> [E1491_D10G132600v 1-233	230	-----V-ADS-----	233
<i>Gossypium_mustelinum</i> [E1491_A10G165200v 1-233	230	-----V-ADS-----	233
<i>Gossypium_mustelinum</i> [E1491_D02G005900v 1-227	224	-----V-ADS-----	227
<i>Gossypium_mustelinum</i> [E1491_A02G005100v 1-227	224	-----V-ADS-----	227
<i>Gossypium_raimondii</i> [B456_011G129400 1-233	230	-----V-ADS-----	233
<i>Gossypium_raimondii</i> [B456_005G005800 1-227	224	-----V-ADS-----	227
<i>Gossypium_tomentosum</i> [E5332_D10G139500v 1-233	230	-----V-ADS-----	233
<i>Gossypium_tomentosum</i> [E5332_A10G178500v 1-233	224	-----V-ADS-----	233
<i>Gossypium_tomentosum</i> [E5332_A02G005200v 1-227	224	-----V-ADS-----	227
<i>Gossypium_tomentosum</i> [E5332_D02G005900v 1-227	224	-----V-ADS-----	227
<i>Theobroma_cacao</i> [TCM_019744 1-228	225	-----V-ADS-----	228
<i>Arabis_alpina</i> [AALP_AA5G141700 1-213	211	-----ADS-----	213
<i>Arabis_nemoralis</i> [ANE_LOCUS23250 1-225	215	-----VVPALAEVADS-----	225
<i>Arabis_nemoralis</i> [ANE_LOCUS15826 1-214	208	-----PELAADS-----	214
<i>Arabis_nemoralis</i> [ANE_LOCUS157979 1-214	208	-----PELAADS-----	214
<i>Brassica_rapa_subsp_pekinensis</i> [WJAH40M 1-215	209	-----LPGL-ADS-----	215
<i>Brassica_rapa_subsp_pekinensis</i> [WJAH40M 1-214	212	-----ADS-----	214
<i>Brassica_rapus</i> [BnaG07g324800 1-216	214	-----ADS-----	216
<i>Brassica_rapus</i> [BnaA03g579600 1-215	209	-----LPGLADS-----	215
<i>Brassica_rapus</i> [BnaCng406600 1-199	192	-----LPAALADS-----	199
<i>Brassica_oleracea_var_oleracea</i> [WJAJ04A03D30C3 1-229	225	-----GT-ADS-----	229
<i>Brassica_oleracea_var_oleracea</i> [WJAJ04A03D30C3 1-216	210	-----LPGLADS-----	216
<i>Brassica_oleracea_var_oleracea</i> [WJAJ04A03D30C3 1-199	192	-----LPAALADS-----	199
<i>Arabidopsis_lyrata_subsp_lyrata</i> [ARALYDRAFT_498889 1-220	210	-----YIPTVEALADS-----	220
<i>Arabidopsis_lyrata_subsp_lyrata</i> [ARALYDRAFT_66800 1-213	211	-----ADS-----	213
<i>Arabidopsis_thaliana</i> [At5g01900 1-217	207	-----VYPAVESLADS-----	217
<i>Arabidopsis_thaliana</i> [At3g15700 1-213	207	-----GPELADS-----	213
<i>Capzella_rubella</i> [CARUB_v10001904eg 1-223	215	-----MESEALADS-----	223
<i>Capzella_rubella</i> [CARUB_v10019623eg 1-213	211	-----ADS-----	213
<i>Eutrema_halophilum</i> [WJJE4MM 1-213	211	-----ADS-----	213
<i>Eutrema_salicigineum</i> [EUTSA_v10010716mg 1-213	208	-----PGSADS-----	213
<i>Eutrema_salicigineum</i> [EUTSA_v10014073mg 1-209	207	-----ADS-----	209
<i>Noccaea_caerulescens</i> [L_TR12690_c0_g1_l1_g_41286 1-219	217	-----ADS-----	219
<i>Noccaea_caerulescens</i> [L_TR12690_c0_g1_l1_g_15690 1-218	216	-----ADS-----	218
<i>Noccaea_caerulescens</i> [GA_TR12421_c0_g1_l1_g_39801 1-218	216	-----ADS-----	218
<i>Noccaea_caerulescens</i> [MP_TR8698_c0_g1_l1_g_27351 1-218	209	-----PYALDTLADS-----	218
<i>Noccaea_caerulescens</i> [MP_TR15565_c0_g1_l1_g_44534 1-213	211	-----ADS-----	213
<i>Noccaea_caerulescens</i> [L_TR17698_c0_g1_l1_g_27127 1-213	211	-----SSV-GT-VPSL-ADA-----	213
<i>Noccaea_caerulescens</i> [GA_TR10503_c0_g1_l1_g_34365 1-213			
<i>Noccaea_caerulescens</i> [L_TR17411_c0_g1_l1_g_56298 1-213	208	-----PGSADS-----	213
<i>Rosa_chinensis</i> [RchQBHb_Chrg3g046096 1-229	218	-----SSL-DI-GSMR-ADS-----	229
<i>Prunus_persica</i> [PRUPE_6G290000 1-253	229	-----SPV-TV-VTVL-SDSKSRQRDGTGMEE-----	253
<i>Prunus_dulcis</i> [ALMOND_3B038990 1-240	229	-----SPV-TV-VTVL-SDS-----	240
<i>Malus_fiondula</i> [DHO4_036312 1-296	260	-----ISNL-SAS-----	296
<i>Malus_baccata</i> [C1H46_04009 1-241	235	-----ISNL-SDS-----	241
<i>Trema_orientale</i> [TorRG33k02_098860 1-233	226	-----EE-----TLR-SDS-----	233
<i>Parasponia_andersonii</i> [PanWU01x14_361630 1-240	233	-----EE-----TLR-SDS-----	240
<i>Rhizophora_mucronata</i> [WJAJ04A2P2644 1-238	236	-----ADS-----	238
<i>Populus_alba</i> [D6086_0000059270 1-240	233	-----S-TMLK-ADS-----	240
<i>Populus_trichocarpa</i> [POPT7_01G133400 1-242	235	-----S-AVLK-ADS-----	242
<i>Populus_trichocarpa</i> [POPT7_006G107300 1-242	235	-----S-TMLT-ADS-----	242
<i>Juglans_regia</i> [LOC10898998 1-249	238	-----SHV-AN-ISLL-SDS-----	249
<i>Juglans_regia</i> [LOC109019257 1-244	233	-----SHS-TK-VSVL-SDS-----	244
<i>Payson_sylvatica</i> [F38_LOCUS102701 1-209	204	-----AML-SDS-----	209
<i>Cucumis_meloni</i> [Cm_4G310980 1-233	222	-----SSV-GT-VPSL-ADA-----	233
<i>Cucumis_melo_var_maluwa</i> [J6676_maffold127G001120 1-249	238	-----SSV-GT-VPSL-ADA-----	249
<i>Cucumis_melo_var_maluwa</i> [J6C27_maffold116G500310 1-233	222	-----SSV-GT-VPSL-ADA-----	233
<i>Cucumis_melo</i> [LOC103502188 1-233	222	-----SSV-GT-VPSL-ADA-----	233

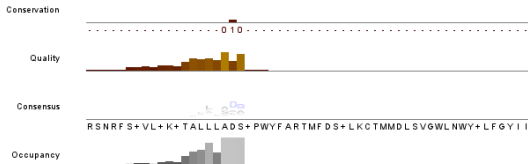
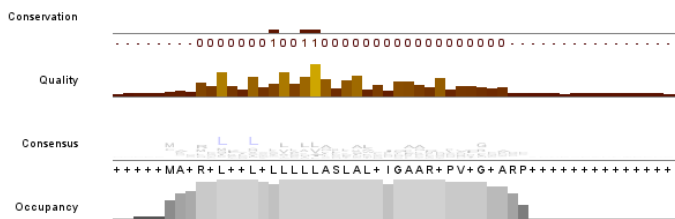


Figure S08. Sequence alignment of PSAPLIPs in angiosperms. Left: Species names with gene ID after the vertical line. If gene ID was not available, protein is was annotated. Conserved sites were shaded with colors in JalViews. Conservation, quality, consensus, and occupancy were calculated and visualized in JalViews by default.

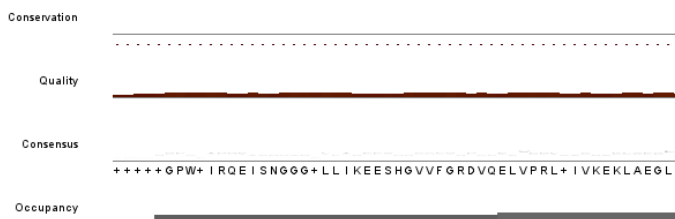
Homo_sapiens[PSAP]prosaposin/1-376
Raphidocelis_subcapitata[Rauc_10640/1-361
Coccomyxa_subellipsoidea/1-332
Chlorella_variabilis[CHLN/CDRAFT_58828/1-332
Chlorella_sorokiniana[C2E21_8413/1-327
Auxenochlorella_protothecoides/1-331
Gonium_pectoralis[GPECTOR_69g440/1-516
Tetrahena_socialis[TSOC_008198/1-430
Chlamydomonas_reinhardtii[CHLRE_05g235700v5/1-429
Chlamydomonas_einhartii[CHLRE_02g105200v5/1-462
Chlamydomonas_eustigma[CEUSTIGMA_g117151/1-345
Klebsomidium_nitens[KFL_001110040/1-305
Chara_braunii[CBR_g3540/1-391
Physcomitrella_patens_subsp_patens[PHYPA_022478/1-333
Physcomitrella_patens_subsp_patens[PHYPA_018982/1-334
Wollemia_nobilis[NA/1-408
Araucaria_cunninghamii[NA/1-406
Picea_sitchensis[NA]A9NUE 1/1-430
Arabidopsis_thaliana[At3g51730/1-213
Arabidopsis_thaliana[At5g01800/1-217

1 NHQKLESNKIPELDMTEVVAPFMANIPLLLYPQDGPRSK----- 40
1 -----MRTLLIAALLAALAPLALASRAAPH----- 25
1 -----MDPQQRCLWGVLCGLTVALAHGAPLQSRSGSDR----- 33
1 -----MRLLLALLG-WVALASAAPVIRSQNNERFRTLRLP----- 33
1 -----MARAAALFCGLLAASVASAPVITSANNPHFRTLRLP----- 35
1 -----MAKRTLAGLVLLVIAGSTMGPLLAQRQTALRGQD----- 34
1 -----MGPRGVLLALLSLVGFASAR-APFMSPVDGPGAN----- 33
1 -----MPAPEGNFGPA----- 11
1 -----MQNMKQLNRAVVPALFSLALLHACGATAARVAVS----- 35
1 -----MPPLKPVLLALVGLFALVNARGSPLVTPPVGDAG----- 34
1 -----MKFVITLALGLFVLARGAYIRDSIRQTP----- 29
1 -----MALALTCIVLLGGVAEGRHMVGPGRPATIQSAP----- 34
1 -----MVGFRRSVLWAVTGLLLLSSVATWLVAVNGRQLAGGDD----- 38
1 -----MAVRNLVWG--VAMLVVCMICVAEGSRVNFQ----- 29
1 -----MAIRNLRWGVAVAVVVLVCMFGIGEGSHVSLG----- 31
1 -----MEMKILSVAFLLIFLSWTSTEARKIVVFER----- 30
1 -----MEMKILSVALLIIFLSWTSTDAKIVVFER----- 30
1 -----MEARMLVIALLVFLITWKNADARKIVIFEHSEIPVGRGFLSHFPSVQ 47
1 -----MGLKAGTFVLLLLGLILV----- 18
1 MGGRFGVLLVFLLSWSCHATNPILLEPFESAHDNDQVCELDCKYVTLVIDYLDQ 55



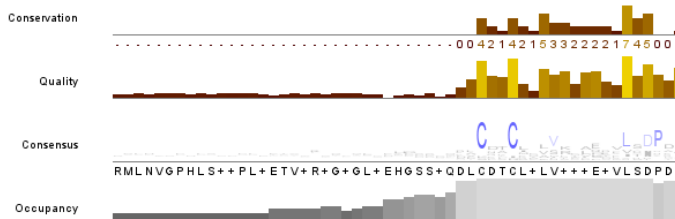
Homo_sapiens[PSAP]prosaposin/1-376
Raphidocelis_subcapitata[Rauc_10640/1-361
Coccomyxa_subellipsoidea/1-332
Chlorella_variabilis[CHLN/CDRAFT_58828/1-332
Chlorella_sorokiniana[C2E21_8413/1-327
Auxenochlorella_protothecoides/1-331
Gonium_pectoralis[GPECTOR_69g440/1-516
Tetrahena_socialis[TSOC_008198/1-430
Chlamydomonas_reinhardtii[CHLRE_05g235700v5/1-429
Chlamydomonas_einhartii[CHLRE_02g105200v5/1-462
Chlamydomonas_eustigma[CEUSTIGMA_g117151/1-345
Klebsomidium_nitens[KFL_001110040/1-305
Chara_braunii[CBR_g3540/1-391
Physcomitrella_patens_subsp_patens[PHYPA_022478/1-333
Physcomitrella_patens_subsp_patens[PHYPA_018982/1-334
Wollemia_nobilis[NA/1-408
Araucaria_cunninghamii[NA/1-406
Picea_sitchensis[NA]A9NUE 1/1-430
Arabidopsis_thaliana[At3g51730/1-213
Arabidopsis_thaliana[At5g01800/1-217

30 -----GIMVTSQKYSFSSSFSS 56
31 -----SGPWAIRQELISDGGGQLFKEESHGTVVFGRDVQELVPRLAIVKEKLAEGWL 81
31 -----SGPWVIRQELISNGGGELLKEESHGTVVFGRDVQELVPRLAIVKEKLAEGWL 81
48 TNCGNGPLLRQEVFNLANGILIRIEEAQRVVFGGDMPESVPLRSRVNQKLAEGLS 102
56 YDNQNELVEALHISCSQIPPLKKQCLSMVDHYTQLFFFTQVSTIKSDQICKRLNLC 110

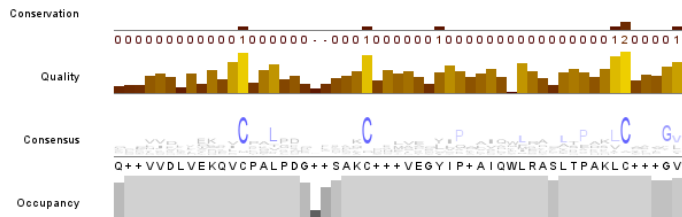


Homo_sapiens[PSAP]prosaposin/1-376
Raphidocelis_subcapitata[Rauc_10640/1-361
Coccomyxa_subellipsoidea/1-332
Chlorella_variabilis[CHLN/CDRAFT_58828/1-332
Chlorella_sorokiniana[C2E21_8413/1-327
Auxenochlorella_protothecoides/1-331
Gonium_pectoralis[GPECTOR_69g440/1-516
Tetrahena_socialis[TSOC_008198/1-430
Chlamydomonas_reinhardtii[CHLRE_05g235700v5/1-429
Chlamydomonas_einhartii[CHLRE_02g105200v5/1-462
Chlamydomonas_eustigma[CEUSTIGMA_g117151/1-345
Klebsomidium_nitens[KFL_001110040/1-305
Chara_braunii[CBR_g3540/1-391
Physcomitrella_patens_subsp_patens[PHYPA_022478/1-333
Physcomitrella_patens_subsp_patens[PHYPA_018982/1-334
Wollemia_nobilis[NA/1-408
Araucaria_cunninghamii[NA/1-406
Picea_sitchensis[NA]A9NUE 1/1-430
Arabidopsis_thaliana[At3g51730/1-213
Arabidopsis_thaliana[At5g01800/1-217

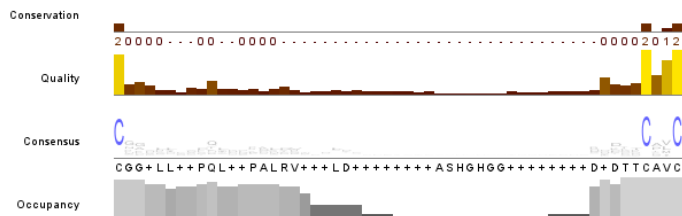
41 -----PQPKDNGDVQDCIQMVTDIQTAVRTNST 69
26 -----DLGCTKDQAVRMKDLMDPFGV 47
34 -----LQSQAPLQSGRWLVKSVKCELSVDVT 60
34 -----EVDNLRSLPQDYQELVTTLESFATDPET 62
36 -----EPAQN-QSPQDYQDVVKTLEDFVSDPT 63
35 -----APACELCEGLVLSLSAYISDPKT 57
34 -----DAQQTQMSVRLLEDPLCDPAA 55
12 -----DAQQALMSVRLLEDPLCDPAA 33
38 -----SAVKEQHVVDRGQDVQDTLLIAMRLLEDALCDDGA 70
35 -----DAQQTLLVVRIVEDLLCDPAA 56
30 -----ESNVGLRAVRVLDNSTITD-TS 50
35 -----LEGTELCEQETLILEAQVVLTDPDN 80
57 SSSDDDDNDVVWKEEGSTAAAAAALPDVHPFKFKGKGVKDFQAEINRPDA 111
32 -----TVARRGKLIQNAFVSRINLQDTQLQSQAELTVLANPDT 69
32 -----TVTRLTKLGRKEAGFAKVNLCQDKQLQSQAQMVLTNPDT 71
82 RMLNVGPHLSGFLPKETVARKGGLTHGSPFQFVQNAQIMVSKQAEVLSNPD 136
82 RMLNVGPHLSGFLPKETVTRKG--LALHGSPFQFVQNAQMEVSKQAEVLSNPD 134
103 KMLNIVPRLSEVPLEKTVSRSG-GLKFGHSSSLGLFNTQMEVSKQAEVLSNPD 156
19 -----SDARSFVDSITSEKVS- 35
111 QAVTPAFASQVHQGNCEACRETVSEVVTCLKDPETKLKIIRLLLECKSKNNYQD 165



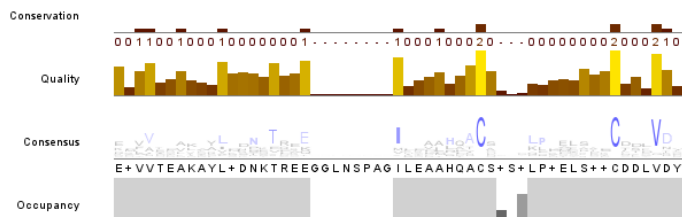
Homo_sapiens[PSAP]prosaposin/1-376
Raphidocelis_subcapitata[Rscu_10640/1-361
Coccomyxa_subellipsoidea/1-332
*Chlorella_variabilis*CHLNCRAFT_58828/1-332
Chlorella_sorokiniana[C2E21_8413/1-327
Auxenochlorella_protothecoides/1-331
Gonium_pectoralis[GPECTOR_69g440/1-516
Tetrahena_socialis[TSOC_008198/1-430
Chlamydomonas_reinhardtii[CHLRE_05g235700v5/1-429
Chlamydomonas_einhartii[CHLRE_02g105200v5/1-462
Chlamydomonas_eustigma[CEUSTIGMA_g11715.1/1-345
Klebsormidium_nitens[KFL_001110040/1-305
Chara_braunii[CBR_g3540/1-391
Physcomitrella_patens_subsp_patens[PHYPA_022478/1-333
Physcomitrella_patens_subsp_patens[PHYPA_018982/1-334
Wollemia_nobilis[WA/1-408
Araucaria_cunninghamii[WA/1-408
Picea_sitchensis[WA]9NUE 1/1-430
Arabidopsis_thaliana[At3g51730/1-213
Arabidopsis_thaliana[At5g01800/1-217



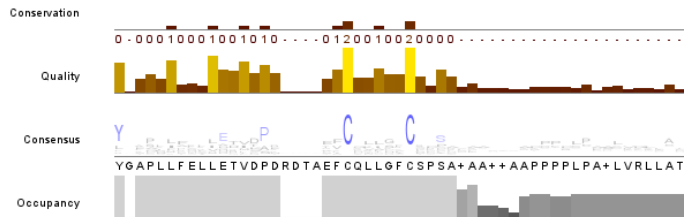
Homo_sapiens[PSAP]prosaposin/1-376
Raphidocelis_subcapitata[Rscu_10640/1-361
Coccomyxa_subellipsoidea/1-332
*Chlorella_variabilis*CHLNCRAFT_58828/1-332
Chlorella_sorokiniana[C2E21_8413/1-327
Auxenochlorella_protothecoides/1-331
Gonium_pectoralis[GPECTOR_69g440/1-516
Tetrahena_socialis[TSOC_008198/1-430
Chlamydomonas_reinhardtii[CHLRE_05g235700v5/1-429
Chlamydomonas_einhartii[CHLRE_02g105200v5/1-462
Chlamydomonas_eustigma[CEUSTIGMA_g11715.1/1-345
Klebsormidium_nitens[KFL_001110040/1-305
Chara_braunii[CBR_g3540/1-391
Physcomitrella_patens_subsp_patens[PHYPA_022478/1-333
Physcomitrella_patens_subsp_patens[PHYPA_018982/1-334
Wollemia_nobilis[WA/1-408
Araucaria_cunninghamii[WA/1-408
Picea_sitchensis[WA]9NUE 1/1-430
Arabidopsis_thaliana[At3g51730/1-213
Arabidopsis_thaliana[At5g01800/1-217



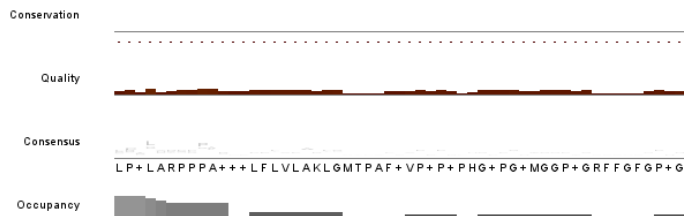
Homo_sapiens[PSAP]prosaposin/1-376
Raphidocelis_subcapitata[Rscu_10640/1-361
Coccomyxa_subellipsoidea/1-332
*Chlorella_variabilis*CHLNCRAFT_58828/1-332
Chlorella_sorokiniana[C2E21_8413/1-327
Auxenochlorella_protothecoides/1-331
Gonium_pectoralis[GPECTOR_69g440/1-516
Tetrahena_socialis[TSOC_008198/1-430
Chlamydomonas_reinhardtii[CHLRE_05g235700v5/1-429
Chlamydomonas_einhartii[CHLRE_02g105200v5/1-462
Chlamydomonas_eustigma[CEUSTIGMA_g11715.1/1-345
Klebsormidium_nitens[KFL_001110040/1-305
Chara_braunii[CBR_g3540/1-391
Physcomitrella_patens_subsp_patens[PHYPA_022478/1-333
Physcomitrella_patens_subsp_patens[PHYPA_018982/1-334
Wollemia_nobilis[WA/1-408
Araucaria_cunninghamii[WA/1-408
Picea_sitchensis[WA]9NUE 1/1-430
Arabidopsis_thaliana[At3g51730/1-213
Arabidopsis_thaliana[At5g01800/1-217



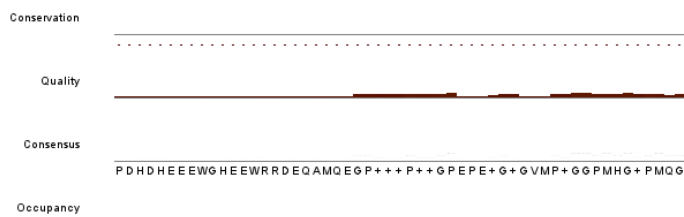
Homo_sapiens[PSAP]prosaposin/1-376
Raphidocelis_subcapitata[Rauc_10640/1-361
Coccomyxa_subellipsoidea/1-332
Chlorella_variabilis[CHLN/CRAFT_58828/1-332
Chlorella_sorokiniana[C2E21_8413/1-327
Auxenochlorella_protothecoidea/1-331
Gonium_pectoralis[GPECTOR_69g440/1-516
Tetrabaena_socialis[TSOC_008198/1-430
Chlamydomonas_reinhardtii[CHLRE_02g235700v5/1-429
Chlamydomonas_einhartii[CHLRE_02g105200v5/1-462
Chlamydomonas_eustigma[CEUSTIGMA_g117151/1-345
Klebsomidium_nitens[KFL_001110040/1-305
Chara_braunii[CBR_g3540/1-391
Physcomitrella_patens_subsp_patens[PHYPA_022478/1-333
Physcomitrella_patens_subsp_patens[PHYPA_018982/1-334
Wollemia_nobilis[NA/1-408
Araucaria_cunninghamii[NA/1-406
Picea_sitichensis[NA]A9NUE1/1-430
Arabidopsis_thaliana[At3g51730/1-213
Arabidopsis_thaliana[At5g01800/1-217



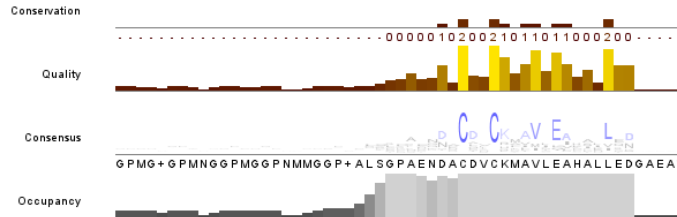
Homo_sapiens[PSAP]prosaposin/1-376
Raphidocelis_subcapitata[Rauc_10640/1-361
Coccomyxa_subellipsoidea/1-332
Chlorella_variabilis[CHLN/CRAFT_58828/1-332
Chlorella_sorokiniana[C2E21_8413/1-327
Auxenochlorella_protothecoidea/1-331
Gonium_pectoralis[GPECTOR_69g440/1-516
Tetrabaena_socialis[TSOC_008198/1-430
Chlamydomonas_reinhardtii[CHLRE_02g235700v5/1-429
Chlamydomonas_einhartii[CHLRE_02g105200v5/1-462
Chlamydomonas_eustigma[CEUSTIGMA_g117151/1-345
Klebsomidium_nitens[KFL_001110040/1-305
Chara_braunii[CBR_g3540/1-391
Physcomitrella_patens_subsp_patens[PHYPA_022478/1-333
Physcomitrella_patens_subsp_patens[PHYPA_018982/1-334
Wollemia_nobilis[NA/1-408
Araucaria_cunninghamii[NA/1-406
Picea_sitichensis[NA]A9NUE1/1-430
Arabidopsis_thaliana[At3g51730/1-213
Arabidopsis_thaliana[At5g01800/1-217



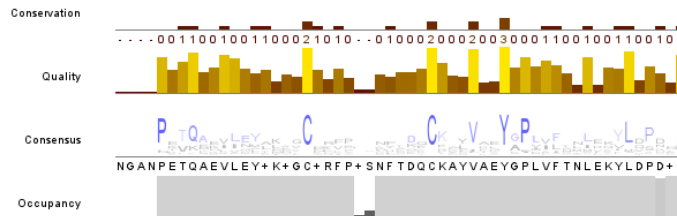
Homo_sapiens[PSAP]prosaposin/1-376
Raphidocelis_subcapitata[Rauc_10640/1-361
Coccomyxa_subellipsoidea/1-332
Chlorella_variabilis[CHLN/CRAFT_58828/1-332
Chlorella_sorokiniana[C2E21_8413/1-327
Auxenochlorella_protothecoidea/1-331
Gonium_pectoralis[GPECTOR_69g440/1-516
Tetrabaena_socialis[TSOC_008198/1-430
Chlamydomonas_reinhardtii[CHLRE_02g235700v5/1-429
Chlamydomonas_einhartii[CHLRE_02g105200v5/1-462
Chlamydomonas_eustigma[CEUSTIGMA_g117151/1-345
Klebsomidium_nitens[KFL_001110040/1-305
Chara_braunii[CBR_g3540/1-391
Physcomitrella_patens_subsp_patens[PHYPA_022478/1-333
Physcomitrella_patens_subsp_patens[PHYPA_018982/1-334
Wollemia_nobilis[NA/1-408
Araucaria_cunninghamii[NA/1-406
Picea_sitichensis[NA]A9NUE1/1-430
Arabidopsis_thaliana[At3g51730/1-213
Arabidopsis_thaliana[At5g01800/1-217



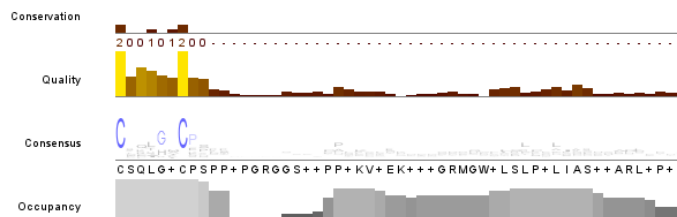
Homo sapiens[PSAP]prosaposin/1-376
Raphidocelis subcapitata[Rscu_10640/1-361
Coccomyxa subellipsoidea/1-332
Chlorella variabilis[CHLN]CDRAFT_58828/1-332
Chlorella sorokiniana[C2E21_8413/1-327
Auxenochlorella protothecoides/1-331
Gonium pectorale[GPECTOR_69g440/1-516
Tetraabaena socialis[TSOC_008198/1-430
Chlamydomonas reinhardtii[CHLRE_05g235700v5/1-429
Chlamydomonas einhardtii[CHLRE_02g105200v5/1-462
Chlamydomonas eustigma[CEUSTIGMA_g117151/1-345
Klebsomidium nitens[KFL_001110040/1-305
Chara braunii[CBR_g3540/1-391
Physcomitrella patens_subsp. patens[PHYPA_022478/1-333
Physcomitrella patens_subsp. patens[PHYPA_018982/1-334
Wollemia nobilis[NA/1-408
Araucaria cunninghamii[NA/1-406
Picea sitchensis[NA]A9NUE/1-430
Arabidopsis thaliana[At3g51730/1-213
Arabidopsis thaliana[At5g01800/1-217



Homo sapiens[PSAP]prosaposin/1-376
Raphidocelis subcapitata[Rscu_10640/1-361
Coccomyxa subellipsoidea/1-332
Chlorella variabilis[CHLN]CDRAFT_58828/1-332
Chlorella sorokiniana[C2E21_8413/1-327
Auxenochlorella protothecoides/1-331
Gonium pectorale[GPECTOR_69g440/1-516
Tetraabaena socialis[TSOC_008198/1-430
Chlamydomonas reinhardtii[CHLRE_05g235700v5/1-429
Chlamydomonas einhardtii[CHLRE_02g105200v5/1-462
Chlamydomonas eustigma[CEUSTIGMA_g117151/1-345
Klebsomidium nitens[KFL_001110040/1-305
Chara braunii[CBR_g3540/1-391
Physcomitrella patens_subsp. patens[PHYPA_022478/1-333
Physcomitrella patens_subsp. patens[PHYPA_018982/1-334
Wollemia nobilis[NA/1-408
Araucaria cunninghamii[NA/1-406
Picea sitchensis[NA]A9NUE/1-430
Arabidopsis thaliana[At3g51730/1-213
Arabidopsis thaliana[At5g01800/1-217



Homo sapiens[PSAP]prosaposin/1-376
Raphidocelis subcapitata[Rscu_10640/1-361
Coccomyxa subellipsoidea/1-332
Chlorella variabilis[CHLN]CDRAFT_58828/1-332
Chlorella sorokiniana[C2E21_8413/1-327
Auxenochlorella protothecoides/1-331
Gonium pectorale[GPECTOR_69g440/1-516
Tetraabaena socialis[TSOC_008198/1-430
Chlamydomonas reinhardtii[CHLRE_05g235700v5/1-429
Chlamydomonas einhardtii[CHLRE_02g105200v5/1-462
Chlamydomonas eustigma[CEUSTIGMA_g117151/1-345
Klebsomidium nitens[KFL_001110040/1-305
Chara braunii[CBR_g3540/1-391
Physcomitrella patens_subsp. patens[PHYPA_022478/1-333
Physcomitrella patens_subsp. patens[PHYPA_018982/1-334
Wollemia nobilis[NA/1-408
Araucaria cunninghamii[NA/1-406
Picea sitchensis[NA]A9NUE/1-430
Arabidopsis thaliana[At3g51730/1-213
Arabidopsis thaliana[At5g01800/1-217



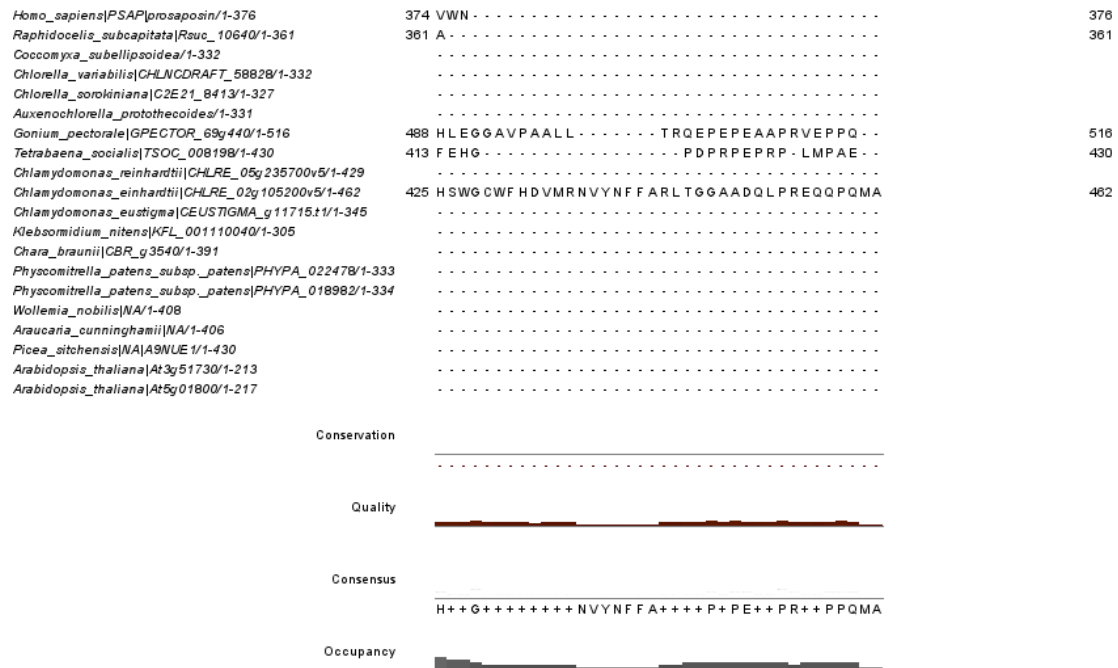


Figure S09. Sequence alignment of PSAPLIPs which contain three SapB-like domains.

Human prosaposin and *Arabidopsis* PSAPLIPs as outliers. Left: Species names with gene ID after the vertical line. If gene ID was not available, protein is was annotated. Conserved sites were shaded with colors in JalViews. Conservation, quality, consensus, and occupancy were calculated and visualized in JalViews by default.

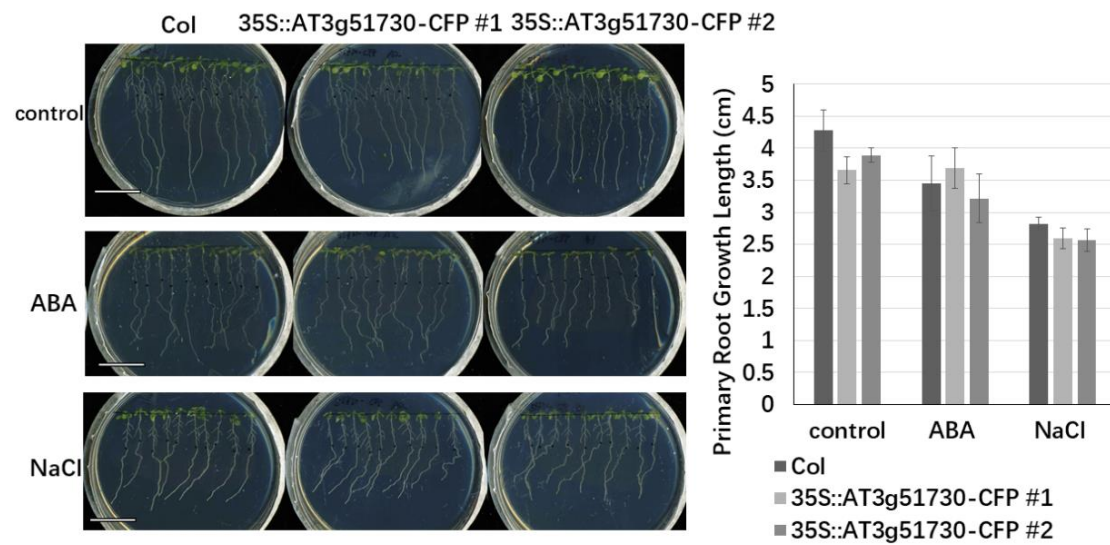


Figure S10. Root growth in *AT3g51730* overexpression plants. 4 DAG seedlings were transferred to media containing ABA (2 μ m) or NaCl (75mM) for another 4 days. Black dots marked the root tip position of 4 DAG seedlings.

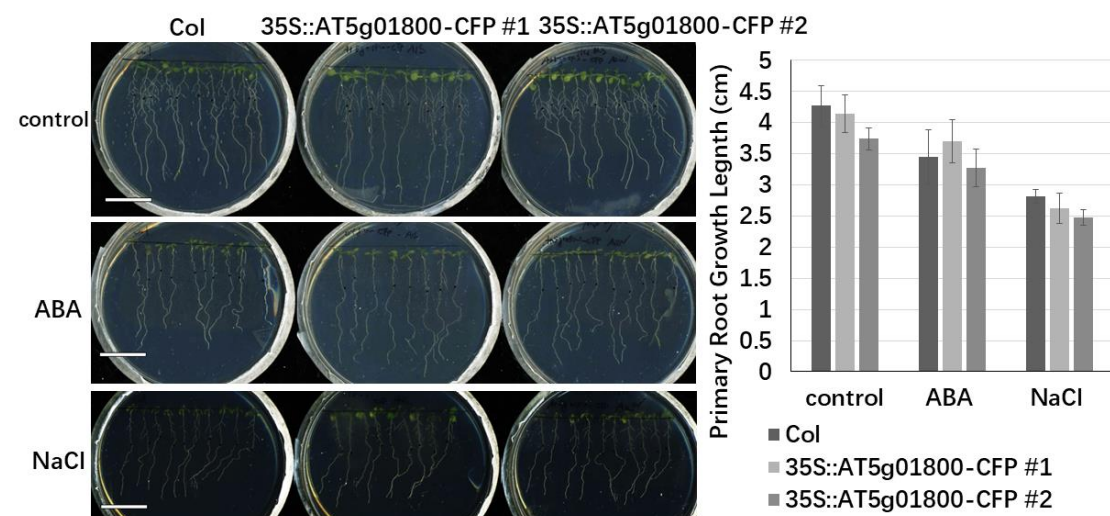


Figure S11. Root growth in *AT5g01800* overexpression plants. 4 DAG seedlings were transferred to media containing ABA (2 μ m) or NaCl (75mM) for another 4 days. Black dots marked the root tip position of 4 DAG seedlings.

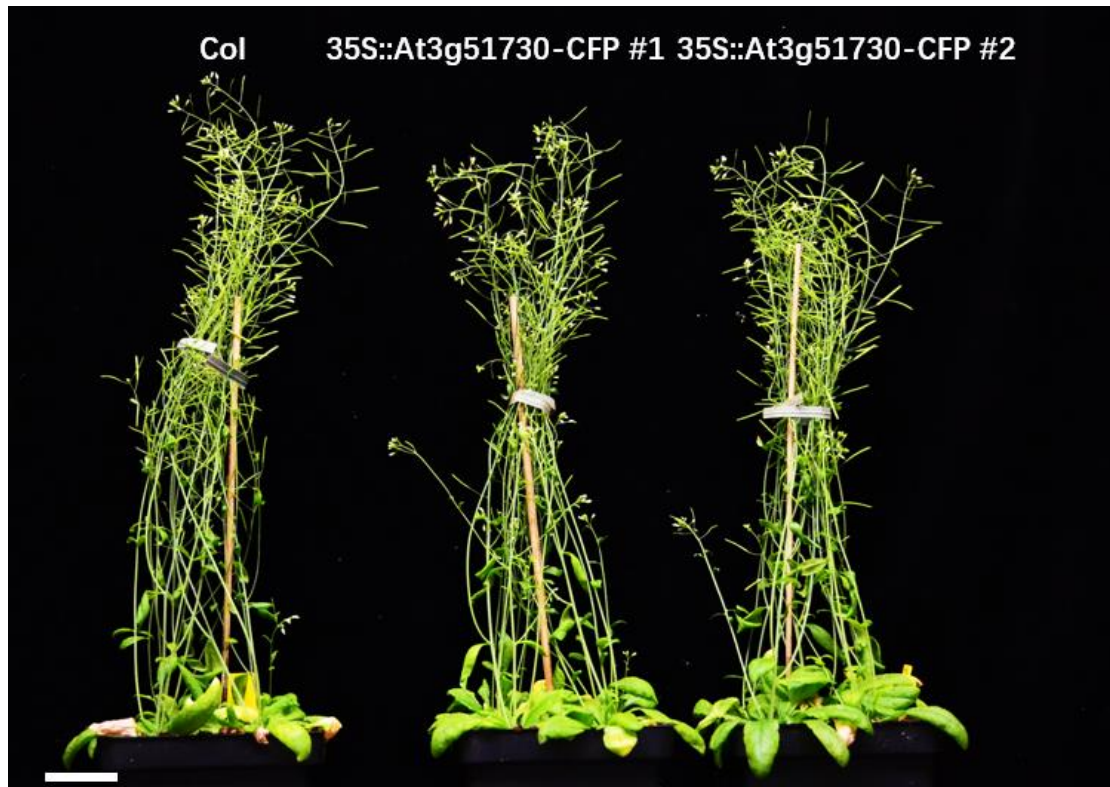


Figure S12. Phenotype of 30 DAG Col and 35S::AtPSAPLIP1-CFP plants. Left: Col; Middle: 35S::AtPSAPLIP1-CFP Line 1; Right: 35S::AtPSAPLIP1-CFP Line 2. Bar=5cm.

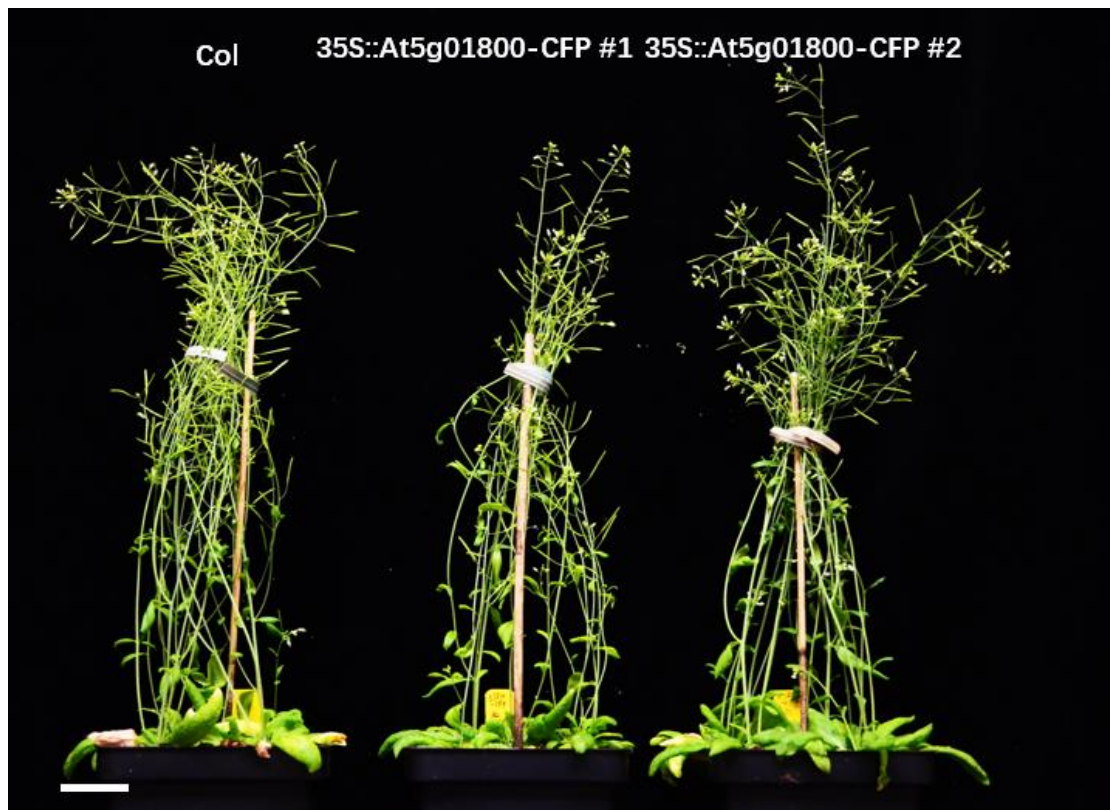


Figure S13. Phenotype of 30 DAG Col and 35S::AtPSAPLIP2-CFP plants. Left: Col; Middle: 35S::AtPSAPLIP2-CFP Line 1; Right: 35S::AtPSAPLIP2-CFP Line 2. Bar=5cm.

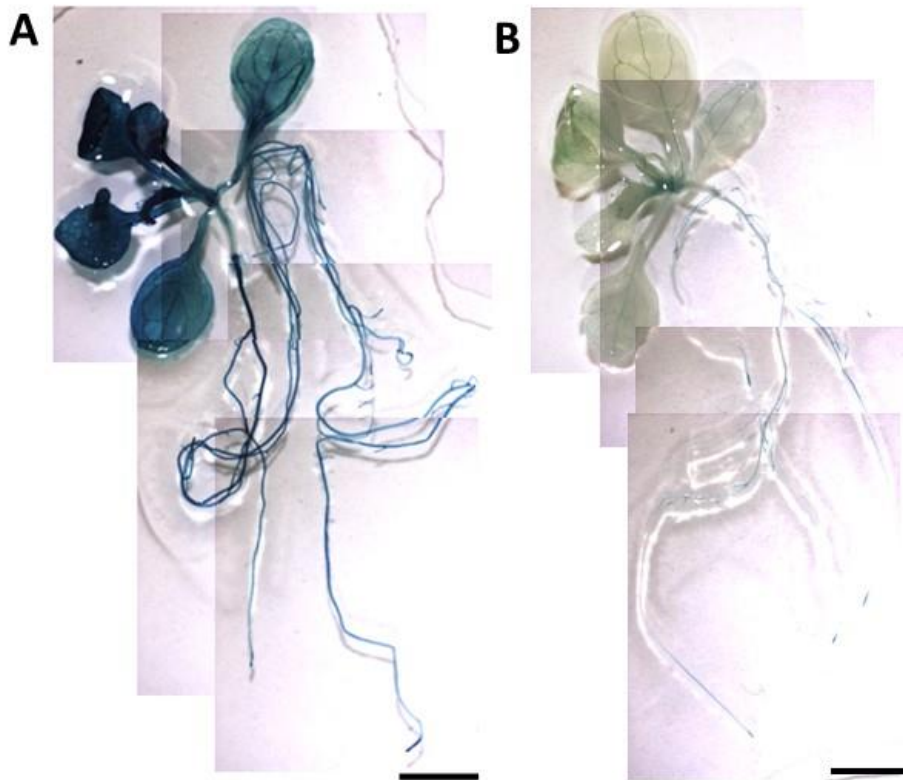


Figure S14. *Arabidopsis* PSAPLIPs promoter::GUS activity in seedlings. (A) PSAPLIP1. (B) PSAPLIP2. 2-week-old seedlings were stained. Bar=0.5cm.

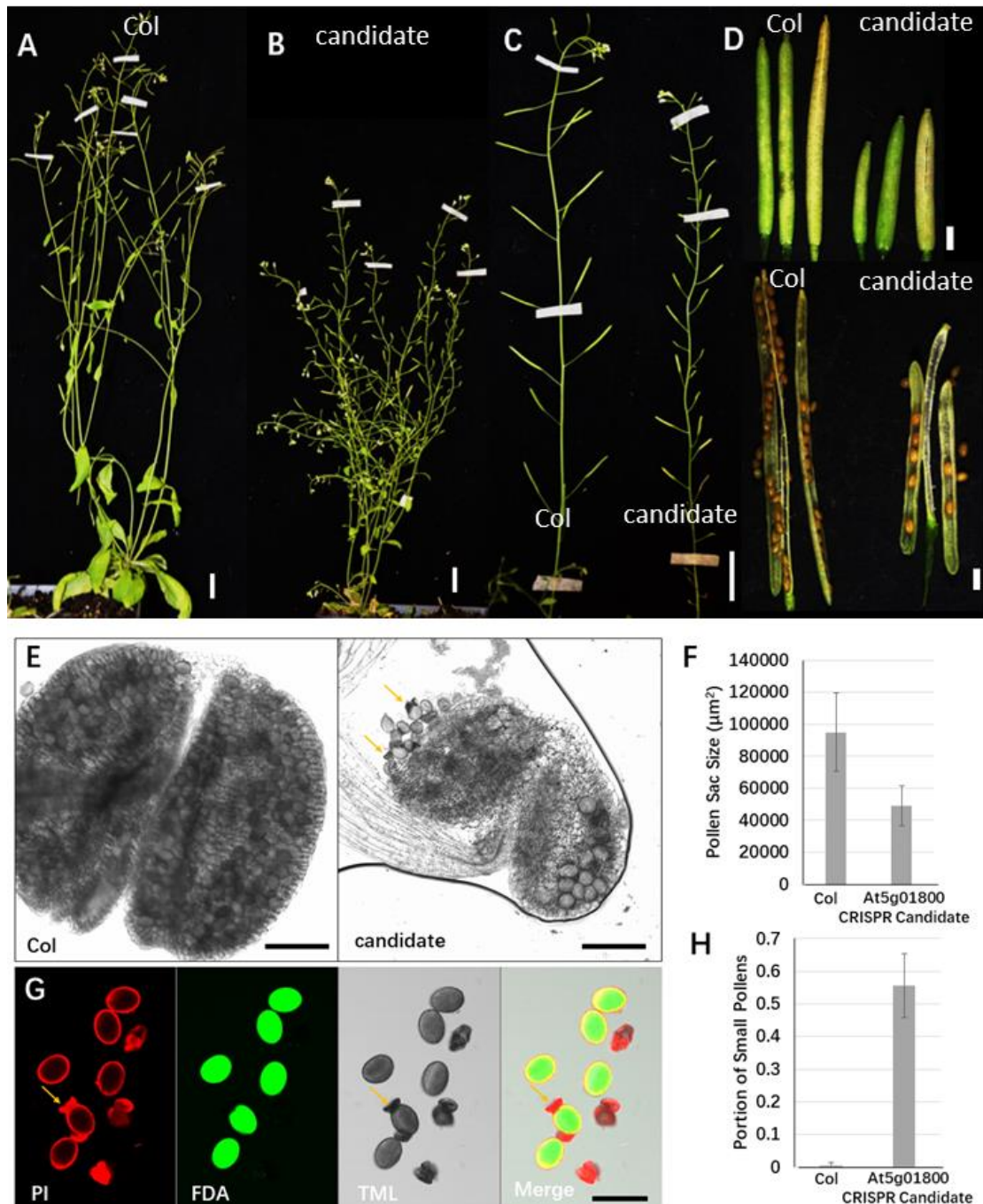


Figure S15. Candidate of *At5g01800* CRISPR mutant. (A) Col plant (B) Possible *At5g01800* CRISPR plant (C) Inflorescence of Col (left) and possible mutant (right). (D) Silique length and seed number in Col (left) and possible mutant (right). (E) DIC image of anther from stage 14 flower in Col (left) and *At5g01800* CRISPR candidate (right). Yellow arrows show the wrinkled pollens. Statistics shown in (F). $P < 0.05$ by Student's *t*-

test (G) propidium iodide (PI)/ fluorescein diacetate (FDA) double staining of the pollens in *At5g01800* CRISPR candidate. Pollens from stage 14 flowers. From left to right: PI, FDA, TML, merge. PI staining indicates dead pollens. (H) Portions of small and wrinkled pollens in Col and *At5g01800* CRISPR candidate. $P < 0.05$ by Student' *t*-test. Bar=2cm in (A)(B)(C), 2mm in (D), 100 μ m in (E), 50 μ m in (G).

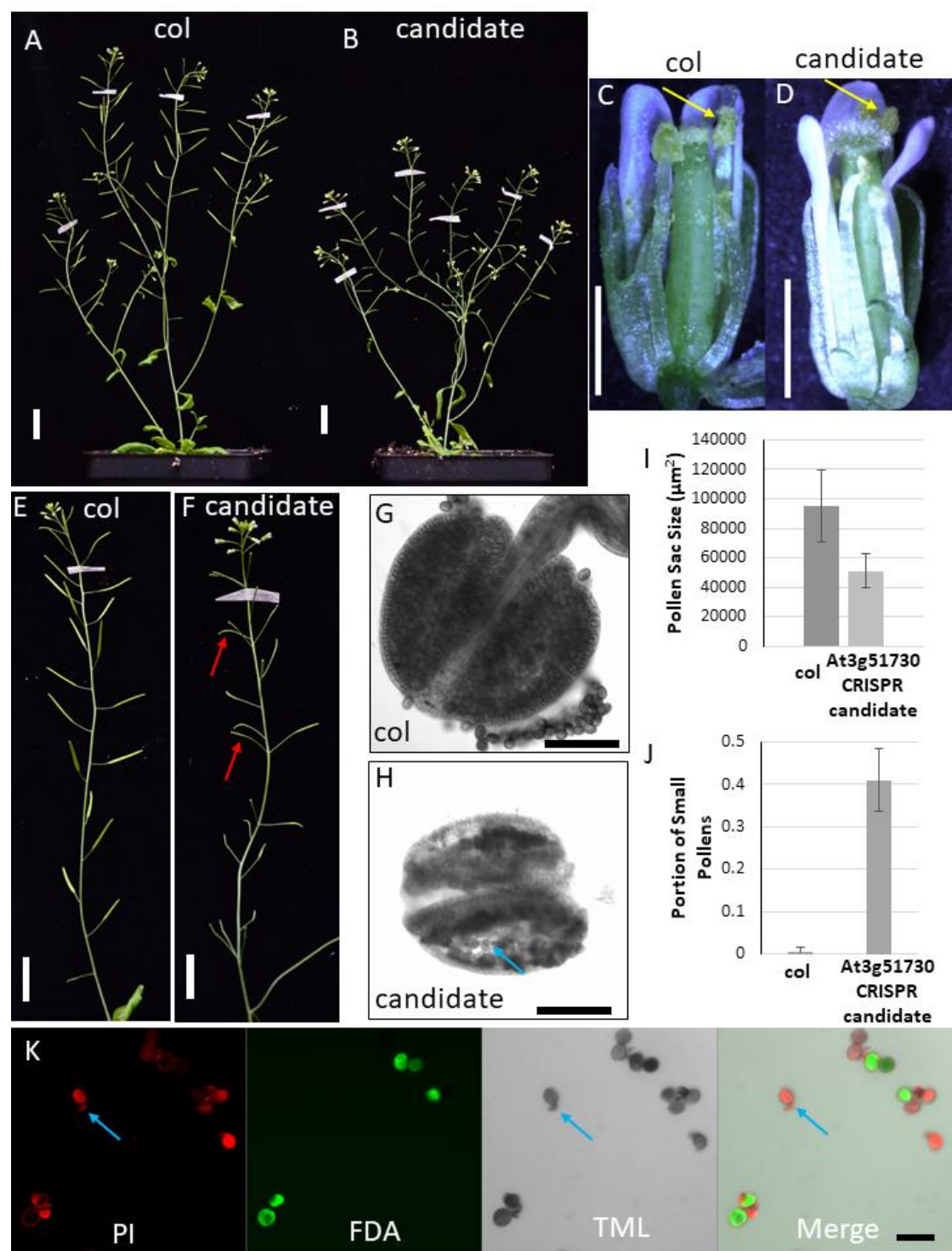
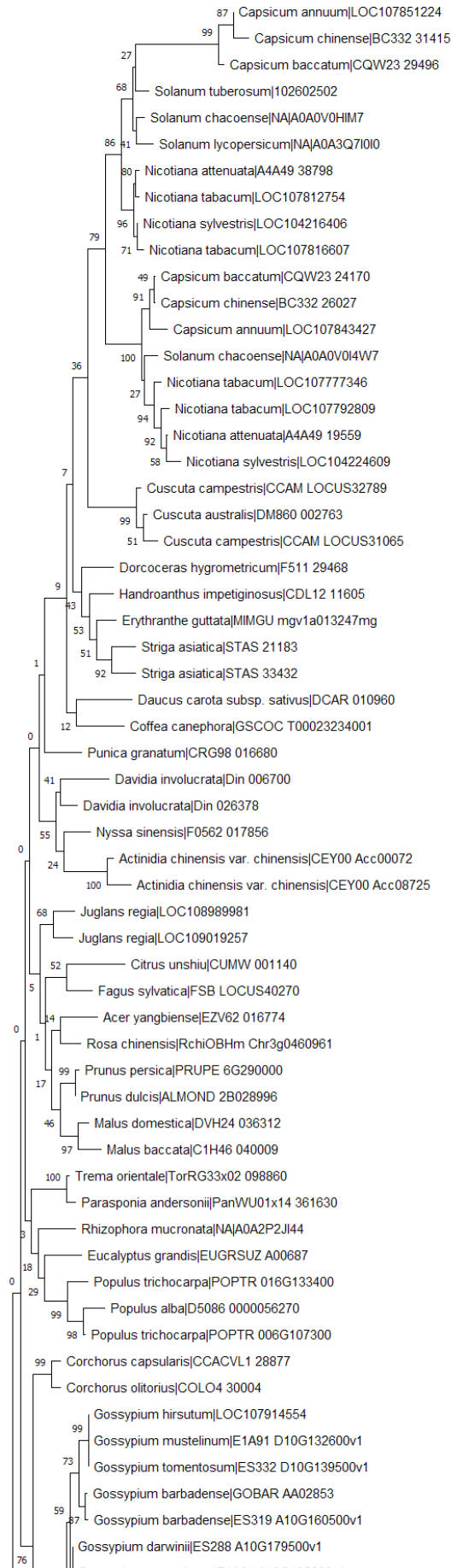
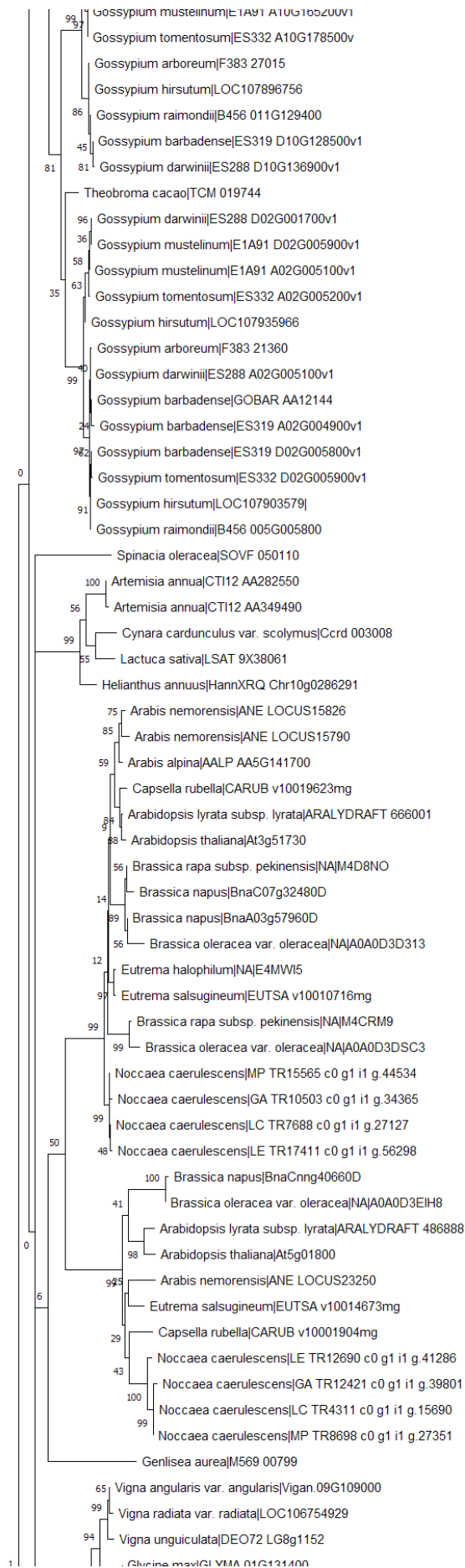


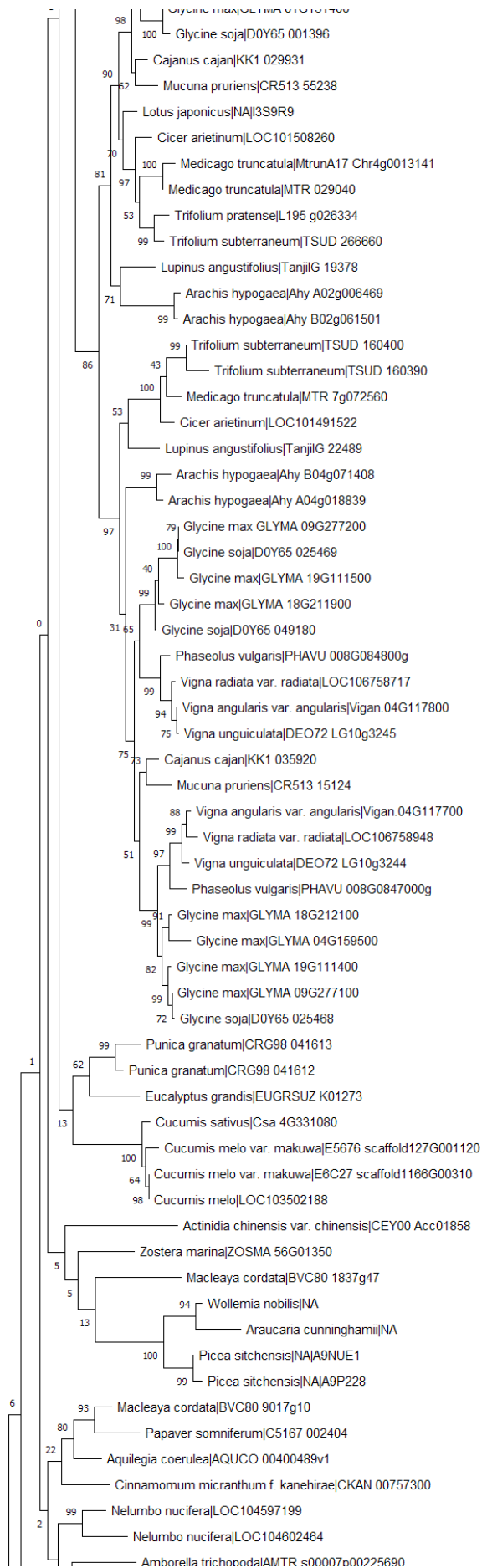
Figure S16. Candidate of *At3g51730* CRISPR mutant. (A) Col plant (B) Possible *At3g51730* CRISPR plant (C) Stage 15 flower of Col and (D) stage 15 flower of CRISPR candidate. Yellow arrows indicate the surface of the anther and released pollens. (E) Inflorescence of Col and (F) Inflorescence possible mutant. Red

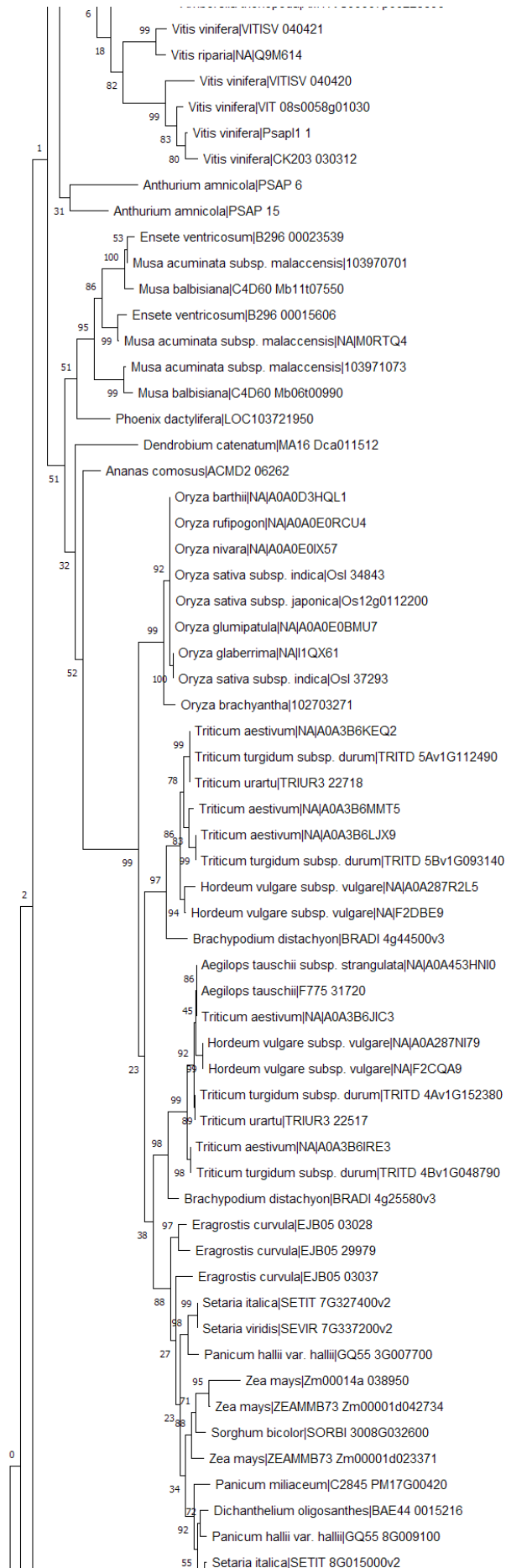
arrows indicate the fertile siliques. (G) DIC image of anther from stage 14 flower in Col and (H) *At3g51730* CRISPR candidate. Blue arrows show the wrinkled pollens. Statistics shown in (I). $P < 0.05$ by Student's *t*-test (J) Portions of small and wrinkled pollens in Col and *At5g01800* CRISPR candidate. $P < 0.05$ by Student's *t*-test. (K) propidium iodide (PI)/fluorescein diacetate (FDA) double staining of the pollens in *At3g51730* CRISPR candidate. Pollens from stage 14 flowers. From left to right: PI, FDA, TML, merge. PI staining indicates dead pollens. Blue arrows show the wrinkled pollens. Bar=2cm in (A)(B)(E)(F), 2mm in (C)(D), 100 μ m in (G)(H), 50 μ m in (K).

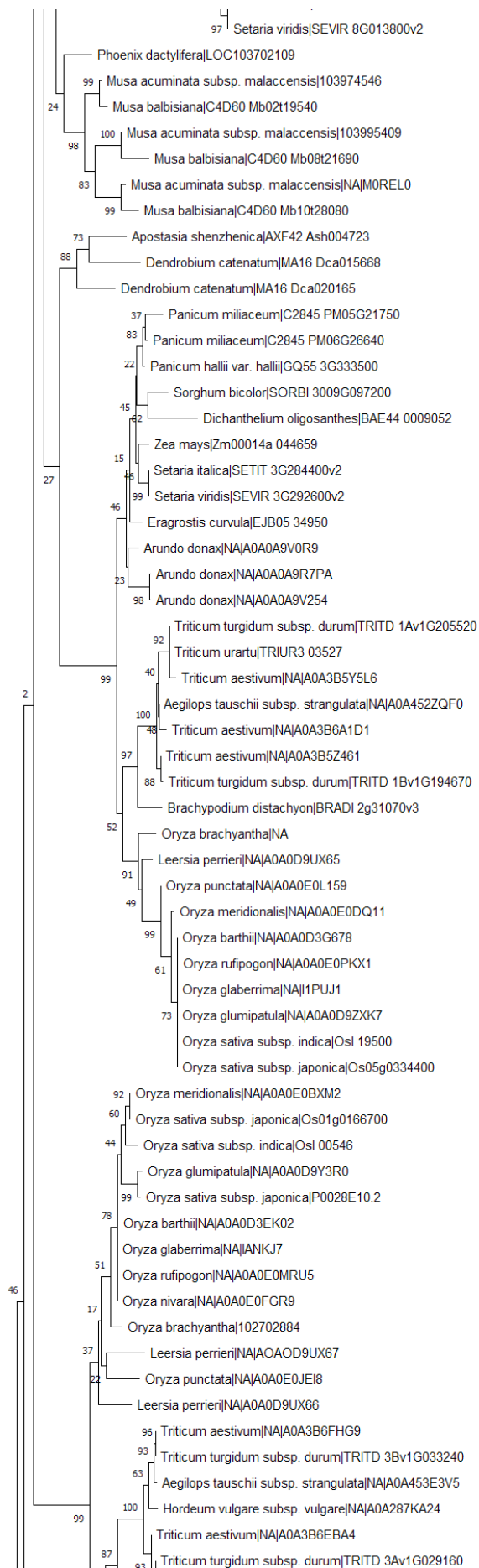
Appendix C Phylogenic tree of plant PSAPLIPs











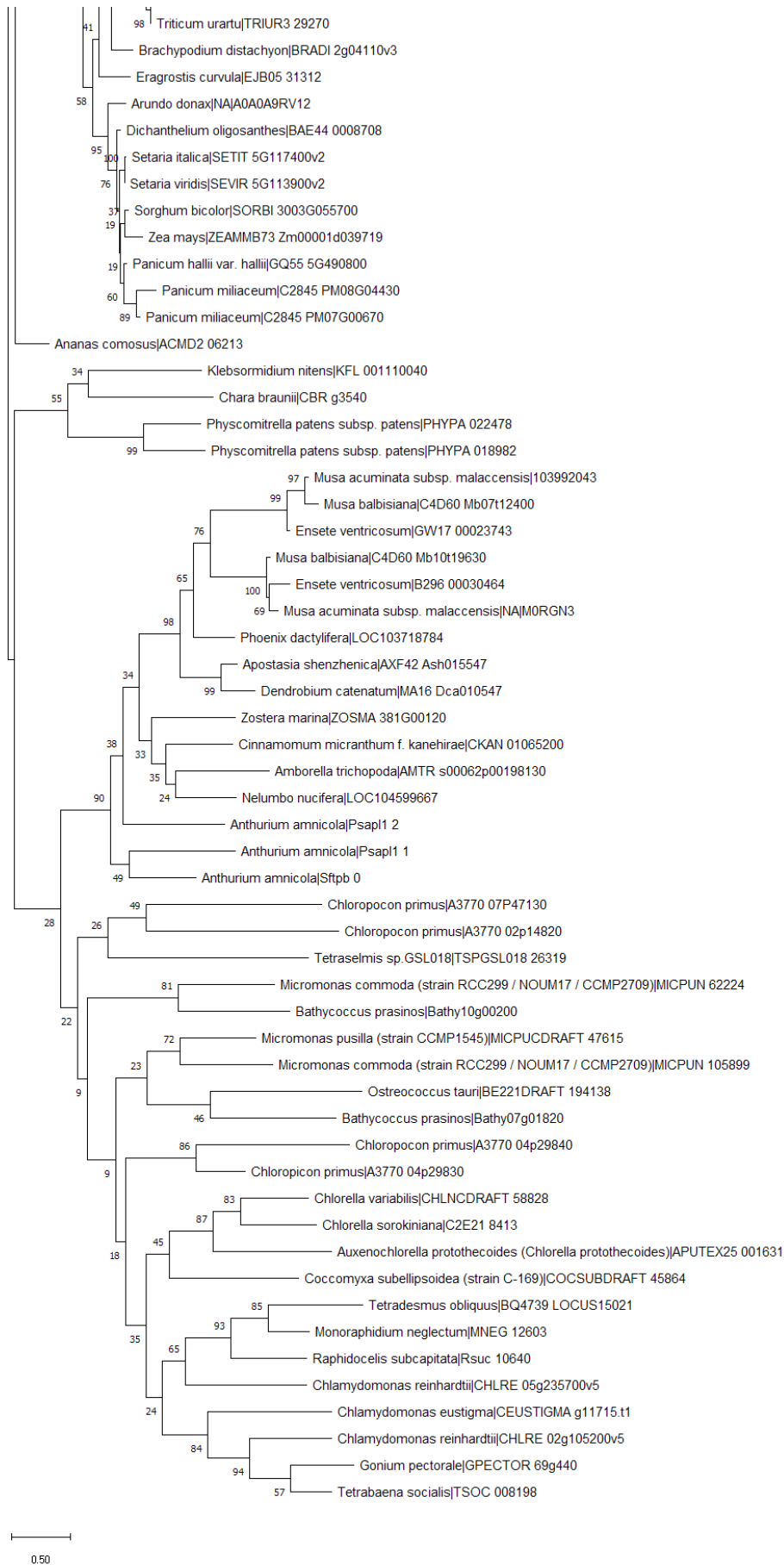


Figure S17. Phylogenetic tree of PSAPLIPs in plants. Phylogenetic tree was constructed in MegaX with maximum likelihood method. Phylogeny test was bootstrap method, with 2000 bootstrap replications. Substitutions type was amino acid with WAG model, which was chosen by Richard et al., 2007. Rates among sites were uniform. All sites were considered. ML heuristic method was nearest neighbor interchange method. No branch swap filter. Number of threads was 3.

Appendix D Materials and Methods

Plant materials

Wild-type *Arabidopsis* plants (*Arabidopsis thaliana* ecotype Columbia-0) and T-DNA insertion mutants were sourced from the Arabidopsis Biological Resource Center, The Ohio State University (ABRC; www.abrc.osu.edu). For germination on soil, seeds were evenly distributed directly on seed germination potting mix. For BASTA selection, BASTA were diluted in sterile water with a concentration of 120mg/ml and sprayed every other day. Green seedlings were selected for further study. For germination on solid 1/4MS media, *Arabidopsis* seeds were surface sterilized by soaking in 20% bleach (containing sodium hypochlorite) for 15 minutes with agitation. Seeds were then rinsed three to five times in sterile water. Seeds were sowed on 1/4MS media supplemented with the appropriate antibiotics and/or chemicals. Seedlings were transferred to soil if needed. Seeds were stratified for 2 days (4°C, dark) then transferred to continuous white light ($100\mu\text{mol m}^{-2} \text{s}^{-1}$) at 22°C conditions for germination.

For germination experiments, *Arabidopsis* seeds (harvested within a month and dried for at least three days) were surface sterilized and plated to 1/4MS. With or without stratification, plates were transferred to continuous white light ($60\mu\text{mol m}^{-2} \text{s}^{-1}$) conditions and assessed every 12 hours for radical emergence.

Nucleic acid isolation

Genomic DNA isolation

Genomic DNA was extracted from *Arabidopsis* seedlings or inflorescence as indicated. Approximately 100mg tissues were collected and grinded in 200µl CTAB extraction buffer and heated at 65°C for at least 30 minutes. Then add 200µl chloroform and isopropanol mix (v: v=24:1) and vortex for 30 seconds. The mixture was centrifuged at 12000rpm for 3 minutes. The supernatant was transferred to new tubes. Add 2µl glycogen (5%) and 167µl isopropanol and mixed thoroughly. The samples were placed at -20°C overnight. Then samples were centrifuged with 12,000rpm at 4°C for 15 minutes. Dispose the supernatant and wash the pellet with 500µl CTAB washing buffer and re-centrifuge with the same setting. Dispose the supernatant, dry the samples with blowing air and dissolve the samples with 500µl sterile water.

CTAB Extraction Buffer

100 mM Tris-HCl (pH 7.5), 25 mM EDTA, 1.5 M NaCl, 2% (w/v) CTAB

CTAB Wash Buffer

70% Ethanol, 10mM Ammonium acetate

Genomic DNA was subsequently used for sequence amplification, or for genotyping by PCR.

Total RNA isolation

Total RNA was isolated from 50-100 mg fresh tissue using ZR plant RNA MiniPrep kit (Zymo Research) according to the manufacturer's instructions. To remove

contaminating genomic DNA, an on-column DNA digest was performed at the time of RNA extraction (DNase I, New England BioLabs) according to manufacturer's instructions.

Plasmid isolation and purification

1.5mL overnight *Escherichia coli* culture ($A_{600}=2-4$) containing the plasmid of interest were pelleted by centrifugation at 10,000 g for 1 minute. Cells were resuspended with 100 μ l solution I. Add 200 μ l solution II and mix well and sit for 1 minute. Mix with 150 μ l solution III and centrifuge at max speed for 15 minutes. Transferred the supernatant to new tubes and add 1ml ethanol. Set the samples on ice for 5 minutes and then centrifuged at 12,000rpm for 15 minutes. Wash the pellet with 1ml 70% ethanol and centrifuge again. Dissolve the pellet with 100 μ l sterile water for further use.

Solution I

50 mM glucose, 10 mM EDTA and 25 mM Tris-HCl, pH 8.0

Solution II

0.2 N NaOH, 1% SDS

Solution III

3 M potassium acetate, 2 M acetic acid

Nucleic acid manipulations

Agarose gel electrophoresis

Nucleic acids were prepared and analyzed using a 1% (w/v) agarose gel made with 1X TAE buffer and molecular grade agarose (Dot Scientific). Nucleic acids were visualized by staining with 0.5% ethidium bromide. Voltage was applied in BioRad submerged horizontal gel. Gels were visualized using Gel Doc™ XR system (BioRad).

50xTAE buffer

Tris base 0.04M, disodium EDTA 0.002M, acetic acid 0.02M.

PCR

For genotyping and colony-PCR, Taq DNA Polymerase with standard Taq buffer (New England BioLabs) was used according to manufacture instructions.

A typical reaction contains components as following:

10x Standard Taq Reaction Buffer 2.5μl

10mM dNTPs 0.5μl

10μM Forward Primer 0.5μl

10μM Reverse Primer 0.5μl

Template variable

Taq DNA Polymerase 0.125μl

Nuclease-Free Water to 25μl

Typical PCR program includes steps as the following:

initial denaturation (95°C) for 3 minutes,

denaturation (95°C), 30s for 35 cycles

annealing (55°C - 60°C), 30s

extension (72°C), 1kb/60s

final extension, 5 minutes (72°C).

final step, 12°C

For promoters, genomic sequence and CDS sequence cloning, Herculase II Fusion DNA polymerase was used.

According to the manufacturer's instructions, a standard reaction contains

5xHerculase II reaction buffer 10µl

dNTP mix (25mM each dNTP) 0.5µl

Template variable

10µM Forward Primer 1.25µl

10µM Reverse Primer 1.25µl

Herculase II DNA polymerase 0.5µl

Nuclease-Free Water to 50µl

Typical PCR program includes steps as the following:

initial denaturation (95°C) for 2 minutes,

denaturation (95°C), 15s

annealing (40°C - 60°C depends on the sequence), 20s

extension (72°C), 1kb/30s

final extension, 3 minutes (72°C).

final step, 12°C

PCR genotyping

PCR genotyping was used to confirm the identity of T-DNA insertion mutants, or putative crosses (F1s). The presence or absence of alleles of interest was determined using diagnostic PCR primer pairs. For known T-DNA insertion mutants, including for F1s, the wild-type allele was identified using primer pairs which spanned the insertion site; the mutant allele was identified using one of the wild-type primers in combination with a T-DNA specific primer (LBb1.3). Homozygous T-DNA insertion mutants were identified by the presence of the allele, and absence of the wild-type allele.

Quantitative RT-PCR

Gene transcript levels were analyzed from total RNA by quantitative RT-PCR using SYBR Green PCR Master Mix kit (Thermo Fisher Scientific) in a BioRad thermo-Cycler, according to manufacturer's instructions. RT-PCR was performed at 95°C for 2 minutes, then 35 cycles of 95°C for 20 s and 54°C for 20 s and 72°C 20s. Melt curve analysis was performed to ensure specificity of the reaction. Threshold values were determined by the CFX manager software (BioRad) and the relative mRNA levels were determined by the $2^{-\Delta\Delta CT}$ method (Pfaffl 2004), using ACTIN 2(ACT2) as a reference gene.

Site-directed mutagenesis PCR

Herculase II was used with the same protocol. After PCR the mixture was digested with DpnI (New England BioLabs) at 37°C for 1 hour. 2µl were used for transformation

into *E.coli*. Plasmids were extracted and sequenced for confirmation.

Purification of PCR products

PCR products were purified using silica. Bands of interest were excised after separation on an agarose gel and two volume: weight ratio of 6M NaI was added. Incubate the agarose in 6M NaI at 55°C for 5-10 min with occasional mixing. Add 10µl of the silica suspension. Vortex gently. Stand for 5 min at room temperature with occasional mixing. One mg of the silica (=10µl of the silica suspension) binds 3-4.5µg of DNA. Spin for 1 min at 12,000rpm. Discard the supernatant and carefully remove residual liquid. Suspend the pellet in 500µl of Solution E. Spin for 1 min at 12000rpm. Discard the supernatant and wash the pellet again. Allow the pellet to air-dry for 10 min. Add an appropriate volume (at least one pellet volume) of sterile water. Vortex gently to resuspend the pellet. Stand for 3 min at 70°C and Spin for 1 min. Transfer the supernatant into a new microfuge tube.

Solution E

50 mM NaCl, 10 mM Tris-HCl pH 7.5, 2.5 mM EDTA, 50%(v/v) ethanol.

Preparation of Silica

Suspend 5 g of silica (Sigma, S-5631) in 50 ml of sterile water. Allow the silica to settle for 2h. Discard the supernatant containing fine particles. Resuspend the pellet with sterile water and re-settle for 2hr. After discarding the supernatant, the packed silica was resuspended in 50ml sterile water to make a final concentration of approximately

100mg/ml.

Digest and ligation reactions

Restriction digests and ligation reactions were carried out as per manufacturer's instructions using 1µL enzyme per 50µL reaction (New England BioLabs). Reactions were incubated overnight at 37°C. Ligation reactions were carried out using T4 ligase (New England BioLabs). Enzymes used for digestion are described in primers.

DNA sequencing

Sequencing of purified DNA was performed by the Eurofins Genomics. Concentrations of primer and purified DNA were as recommended by Eurofins Genomics. Sequence analysis was performed using the VectorNTI® software (Life TechnologiesTM).

Gateway Cloning

Gateway Cloning Binary vectors for in planta genetic modification were constructed using Gateway technology (InvitrogenTM) as follows.

TOPO reaction

TOPO of entry clones Gateway® compatible entry clones are prepared according to manufacturer's instructions. TOPO-compatible overhang was incorporated into the

fragment of interest during PCR amplification by including a CACC at the beginning of the forward primer. Mix the following components for reaction:

Fresh PCR product 0.5–4 μ l,

Salt Solution 1 μ l,

Water add to a total volume of 5 μ l,

pENTR/D-TOPO® vector 1 μ l.

Final Volume 6 μ l.

The mixture was incubated at least 30 minutes at room temperature. 2 μ l of this reaction was transformed into Mach1-T1 (Invitrogen™) chemically competent *E. coli* cells. Kanamycin was added to the media for selection. Positive clones were identified by colony PCR, in which a small amount of bacterial colony was incorporated directly into a PCR reaction. The plasmids were sequenced using M13F and M13R primers and additional internal primers where necessary.

BP reaction

BP entry clones are prepared according to manufacturer's instructions. BP-compatible overhang was incorporated into the fragment of interest during PCR amplification by including attB1 and attB2 at the beginning of the forward and reverse primers. Mix the following components for reaction:

Fresh PCR products 1 μ l,

BP Clonase II enzyme mix 0.5 μ l,

pDONR/Zeo vector 1 μ L,

water add to a total volume 10 μ L.

The mixture is incubated at least 3 hours at room temperature. 4 μ L of this reaction was transformed into Mach1-T1 (InvitrogenTM) chemically competent *E. coli* cells. Zeocin was added to media for selection. Positive clones were identified by colony PCR. The plasmids were sequenced using M13F and M13R primers and additional internal primers where necessary.

LR reactions

Gateway destination vectors contain attR Gateway compatible sites flanking a ccdB death gene. These plasmids were cultured in DB3.1 ccdB survival *E. coli* cells (InvitrogenTM). Destination vectors used in this project included the pEARLEYGATE102 (pEG; Earley et al. 2006), the pUBC series (Curtis and Grossniklaus 2003), pH7WGC2, pH7WGR2, and pGBW3 (for promoter analysis, Karimi et al. 2002). Expression clones were prepared using LR Clonase II Gateway kit (InvitrogenTM) according to the manufacturer's instructions. Molar ratios were carefully balanced. LR reactions were incubated for 3 hours at room temperature. 2 μ l was transformed into *E. coli* cells. Proper antibiotics are added in the media for selection depending on the destination vectors. Positive clones were identified by colony PCR.

Transformation of bacteria

Mach1-T1 chemically competent *E. coli* cells were prepared by Mix and Go! Transformation Buffer Set (Zymo Research) according to the manufacturer's instruction. Briefly, *E. coli* cells were thawed on ice, then incubated on ice with 2 µl of plasmids cloning mixture for 30 minutes. 300 µl room temperature LB was added. Transformed cells were shaken horizontally at 37°C, 200rpm for 1 hour, then inoculated onto LB agar plates containing the relevant antibiotic and incubated at 37°C overnight.

For transformation of *Arabidopsis*, vectors of interest were transformed into electrocompetent *Agrobacterium* cells by electroporation. GV3101 *agrobacterium* cells were thawed on ice, then incubated on ice with 1µl plasmid DNA for 30 minutes. Cells were transferred to a chilled 1 mm electroporation cuvette and electroporated using a Bio-Rad Micropulser® ("AGR" settings). Transformed cells were shaken horizontally at 28°C, 200rpm for 3 hours, then inoculated onto LB agar plates containing the relevant antibiotics and incubated at 28°C for two days.

Transformation of *Arabidopsis* plants by floral dipping

A modified version of the floral dip method (Clough and Bent 1998) was used for *agrobacterium* -mediated transformation of *Arabidopsis* plants. Briefly, *agrobacterium* carrying the desired construct was streaked to LB agar plates (with selection) and incubated at 28°C for 24 hours. *Agrobacterium* was resuspended in 80 ml fresh LB media to an OD600 of 2-2.5.

The bacteria were harvested by centrifuging with 3000g at room temperature and then suspended in 5% sucrose solution containing 0.02% Silvet-77. *Arabidopsis* floral buds were dipped in this solution and wrapped with plastic wraps to keep the humidity. The plants were covered in large plastic bags in the dark overnight. Positive transformants were identified by germinating on agar plates supplemented with hygromycin, or by germinating on soil and treating seedlings with a selective herbicide BASTA. Then selected seedlings were confirmed by LSM confocal microscope

Crossing *Arabidopsis* genetic lines

Suitable inflorescences containing healthy flower clusters were chosen. Elongating siliques and open flower buds were removed under a dissecting microscope. Ideal flower buds (large, but without an exposed stigma) were carefully emasculated, avoiding damage to the stigma, style petals and sepals. The other flower buds were removed to avoid confusion. The emasculated flowers were ready after 24 hours for pollination. Mature anthers from the paternal parent were collected and used to spread pollens onto the exposed stigmatic region. Cross-pollinated flowers were labeled in a piece of paper tape. Successful crosses were identified in the F2 generation.

Protein assays

Total protein extraction from imbibed seeds

Protein extraction

300mg *Arabidopsis* tissues were ground with a grind stick in Eppendorf tubes with liquid nitrogen. The ground tissues were resuspended in 300 μ L protein extraction buffer (50 mM sodium citrate, pH 5.5; 5% SDS (w/v); 0.01% BSA (w/v); 150 mM NaCl; 2% (v/v) β -mercaptoethanol and 1 μ L of protease inhibitor cocktail (Genesee Scientific). The mixture was incubated for 60 minutes at 100° C. Samples were centrifuged at 4° C, 14,500g for 30 minutes and the supernatant was collected. The samples were stored in -80° C if not used immediately.

Glycosylation test

Glycosylation was detected by Endo Hf (New England BioLabs) digestion according to manufacturer's instruction. Briefly, 17 μ L extracted protein sample was added with 2 μ L 10xGlycoBuffer 3, 1 μ L Endo Hf. The samples were incubated at 37°C for 1 hour. Then the sample was used for SDS-PAGE and Western blot.

SDS-PAGE

Total proteins were separated by SDS polyacrylamide gel electrophoresis (SDS-PAGE). 10 μ L samples were prepared by adding 2 μ L of 6X SDS (sodium dodecyl sulfate) loading buffer (1.2g SDS, 0.01% bromophenol blue, 4.7ml glycerol, 1.2ml Tris 0.5M pH=6.8, 2.1ml water). Samples were loaded onto 12% polyacrylamide 0.75mm 10-well or 15-well gel (Bio-Rad®). Precision Plus Protein Dual Color Standards (Bio Rad) was used to mark band size. Electrophoresis was carried out in 1X Running Buffer (3g of

Tris base, 14.4g of glycine, and 1g of SDS in 1000 ml water) at 120V for approximately 4 hours or until the dye front reached the front of the gel.

Western blot

For immunoblotting, proteins were transferred to polyvinylidene difluoride (PVDF) membrane in Tris-glycine-methanol transfer buffer (2.9g glycine, 5.8g Tris, 0.37g SDS 100mL methanol, 900mL water) at 120V for 80 minutes at 4°C and then rinsed briefly in 1xPBS. Membranes were blocked overnight at 4°C in blocking buffer (5% non-fat milk in 1xPBS with 0.02% Tween20) or 1.5 hours at room temperature. The membrane was rinsed gently with washing buffer (1% non-fat milk in 1x PBS with 0.02% Tween20) for three times each for 15 minutes. The membrane then was incubated with primary antibody (anti-HA) in blocking buffer overnight at 4°C or 1.5 hours at room temperature. The membrane was rinsed with washing buffer for three times each for 15 minutes. Then the membrane was incubated with secondary antibody (anti-rabbit digoxigenin) at room temperature for 1.5 hours. The membrane was rinsed with washing buffer for three times and each time for 15 minutes. Proteins were visualized using a SuperSignal West Femto Kit (Thermo Scientific). Images were taken by C-DiGit Blot Scanner (LI-COR).

Coomassie blue staining

For visualization of seed storage proteins, the gel stained by incubating overnight

in 20ml Coomassie staining solution (0.1% Coomassie bright blue in 50% methanol, 10% acetic acid). The gel was de-stained for 3 hours with de-staining solution (10% acetic acid, 50% methanol). At least two changes of this solution until the background was nearly clear.

Histochemistry

Promoter GUS activity was visualized in planta using a GUS or histone10 2A (H2A) tagged with fluorescent protein reporter system. Promoter fragments were amplified from wild type genomic DNA. Promoters were inserted upstream of β -glucuronidase (GUS) in the pGBW3 destination vector for transformation into wild type *Arabidopsis*.

For *ASPAs*, target promoters were replaced for the UBQ10 promoter in pUBC::YFP-Dest or pUBC::mCherry-Dest vectors. Then H2A was incorporated by LR reaction. For all constructs, putative transformants were identified by hygromycin (GUS constructs) or BASTA (H2A-YFP/mCherry construct) and confirmed by genomic DNA PCR using promoter-specific forward GUS reverse primers.

For GUS staining detection, plant tissues were fixed in cold 90% acetone for 30 minutes, then washed twice in GUS buffer before staining. Samples were infiltrated with GUS buffer under vacuum for 10 minutes, then incubated at 37°C for 48 hours. Tissue was cleared in 70% ethanol overnight and repeated several times until the tissue becomes clean and clear. The sample was mounted on microscope slides for visualization.

GUS staining buffer

Sodium phosphate buffer (pH=7) 100mM, EDTA 10mM, Triton X-100 (w/v) 0.1%, potassium ferrocyanide 2mM, potassium ferricyanide 2mM, X-glucuronide 0.5mg/ml.

Microscopy and imaging

Microscopy

Confocal microscopy was carried out using a Zeiss LSM 710 Confocal laser scanning microscope (Carl Zeiss, Germany) with Axio Imager 2. Pixel dwell time was 0.01 ms. The master gain was always set to less than 893, with a digital gain of 1.5. For RFP/mCherry acquisition: 594 nm (5%) excitation and 588-696 nm emission. For YFP acquisition: 514 nm (5%) excitation and 519-560 nm emission. For GFP: 488 nm (5%) excitation and 493-598 nm emission. For CFP: 458 nm (5%) excitation and 453-580 nm emission. For PI: 543 nm (5%) excitation and 583-718 nm emission. For FDA: 488 nm (5%) excitation and 493-583 nm emission. Quantification of fluorescence intensity was analyzed using ZEN Lite 2012.

Image production

Post-processing of microscopy images was performed using Fuji/ImageJ and associated plugins ([www. http://fiji.sc/](http://fiji.sc/); Schneider et al. 2012; Schindelin et al. 2012), or Zeiss ZEN Black v10.0 (Carl Zeiss, Germany; <http://www.zeiss.com/microscopy/>). Image quantification was carried out using ImageJ.

Bioinformatics

Primary and Secondary Structure Prediction

Hydropathy plot was drawn in ExPASy with Kyte and Doolittle method. Window size was 9 with the linear weight variation model. Structure prediction was conducted in Phyre2. Each SapB-like domain was predicted separately. Predicted structure of AtPSAPLIP1 and AtPSAPLIP2. Final images were visualized with EzMol.

Sequence Alignment

PSAPLIPs protein sequences were selected in EggNOG (<http://eggnog5.embl.de/>) and Uniprot (www.uniprot.org). In EggNOG, sequences were identified via pairwise ortholog predictions with *AT3G51730*. 167 sequences from 67 species were outputs. In Uniprot, sequences were screened by searching keyword saposin. Only sequences in Viridiplantae were chosen for further screening. The sequences which were annotated as fragments were removed. Aspartic proteases were removed as well. For those sequences without the gene ID, if sequences similarity was above 95%, the longer one was kept. If the annotated SapB-like domain length was below 50 amino acid residues, the corresponding sequences were also removed.

After first try of alignment, the sequences belonging to the neucleophosmin family were removed. The remaining sequences were considered valid PSAPLIP proteins in plants and for further analysis.

Alignment was conducted in MegaX with Clustal MUSCLE method. The parameters were as following: gap open -2.9, gap extend 0, hydrophobicity multiplier 1.2, max memory in MB 2048, max iterations 16, cluster method UPGMA, cluster method UPGMA, min diag length 24. Some manual adjustments were applied for gap positions for better alignments.

To search for conservative positions, the sequence that only contain one SapB-like domains were removed because they may be incomplete sequences if there are errors in predictions. Sequences in green algae, liverworts, mosses and gymnosperms were aligned separately due to their variable number of copies of SapB-like domains. Human prosaposin and Arabidopsis PSAPLIPs were chosen as the outlier.

Images were processed with JalView. Color was added by Taylor method with conservation level 85%. Annotation was calculated automatically.

Phylogenetic tree construction

Phylogenetic tree of plant PSAPLIPs were constructed in MegaX with maximum likelihood method. Phylogeny test was bootstrap method, with 2000 bootstrap replications. Substitutions type was amino acid with WAG model. Rates among sites were uniform. All sites were considered. ML heuristic method was nearest neighbor interchange method. No branch swap filter. Number of threads was 3.

Statistical analysis

All means and standard errors were calculated using Microsoft Excel 2013. Where indicated, statistical significance was determined using a Student's *t*-test, with tails=2 and type=3 (independent samples of unequal variance; Microsoft Excel 2013) unless otherwise indicated. Pearson's chi square analyses were performed to determine the segregation ratios for single insertion segregation where mentioned.

Appendix E Supplemental Tables

Table S01. Primer List in this dissertation.

Primer Name	Direction	Use	Sequence
<i>aspa2-1</i>	Forward	Genotyping for T-DNA	TTTTTGGAGCATTATTGCGAC

SALK_097505 LP		insertion	
<i>aspa2-1</i> SALK_097505 RP	Reverse	Genotyping for T-DNA insertion	AATTCGAATGTGTGACAAATCG
<i>aspa2-2</i> SALK_021601 LP	Forward	Genotyping for T-DNA insertion	TTTTTGGAGCATTATTGCGAC
<i>aspa2-2</i> SALK_021601 RP	Reverse	Genotyping for T-DNA insertion	ATTGATCCTGAGCCGTAATGG
<i>aspa1-1</i> SALK_092586 LP	Forward	Genotyping for T-DNA insertion	GTCTTGGTGCAATTGAGATT
<i>aspa1-1</i> SALK_092586 RP	Reverse	Genotyping for T-DNA insertion	AATAGCATTTTGATGATGGC
<i>aspa1-2</i> SALK_041027 LP	Forward	Genotyping for T-DNA insertion	ATGAAGATATACTCTAGAAC
<i>aspa1-2</i> SALK_041027 RP	Reverse	Genotyping for T-DNA insertion	ATACCAAACAGGAGCAGCTT
<i>aspa3-1</i> CS330614 RP	Forward	Genotyping for T-DNA insertion	ATGGGAACTAGGTTCCAATC
<i>aspa3-1</i> CS330614 LP	Reverse	Genotyping for T-DNA insertion	ACATCATCATTGCTAAAGTA
<i>aspa3-2</i> SK36621 LP	Forward	Genotyping for T-DNA insertion	CTATTTGGATGCTCAATACT

<i>aspa3-2</i> SK36621 RP	Reverse	Genotyping for T-DNA insertion	GGAGATCACCCATGTCAAAC
<i>aspa3-3</i> SALK_056711C LP	Forward	Genotyping for T-DNA insertion	CATAAAGGTTACTGGCAGTT
<i>aspa3-3</i> SALK_056711C RP	Reverse	Genotyping for T-DNA insertion	TACTGCAGACAGACATGAAT
LBb1.3	Forward	Genotyping for T-DNA insertion	ATTTTGCCGATTTCGGAAC
ASPA2 qRT F	Forward	Quantitative PCR	TTGAGGCAGAACATGACTCA
ASPA2 qRT R	Reverse	Quantitative PCR	CACGGCTTCTGCGAAGCCAA
ASPA1 qRT F	Forward	Quantitative PCR	GAGCGCATATTGAACTACGT
ASPA1 qRT R	Reverse	Quantitative PCR	GGCTGCCTCTGCAAACCCGA
ASPA3 qRT F	Forward	Quantitative PCR	ACACAAGAACGCATACTCGC
ASPA3 qRT R	Reverse	Quantitative PCR	AGCAGCTTTGGCGAATCCAA
ACT2 qRT F	Forward	Quantitative PCR	ACACTGTGCCAATCTACGAGGGT T
ACT2 qRT R	Reverse	Quantitative PCR	ACAATTTCCCGCTCTGCTGTTGTG
ASPA2 F attB1	Forward	CDS cloning	GGGGACAAGTTTGTACAAAAAA GCAGGCTCCATGTCCCCTATAGAT CC
ASPA2 R NS attB2	Reverse	CDS cloning	GGGGACCACTTTGTACAAGAAA

			GCTGGGTCCACGGCTTCTGCGAA GCCAA
ASPA1 F attB1	Forward	CDS cloning	GGGGACAAGTTTGTACAAAAA GCAGGCTCCATGAAGATATACTCT AGAAC
ASPA1 R ns attB2	Reverse	CDS cloning	GGGGACCACTTTGTACAAGAAA GCTGGGTCGGCTGCCTCTGCAAA CCCGA
ASPA3 F attB1	Forward	CDS cloning	GGGGACAAGTTTGTACAAAAA GCAGGCTCCATGGGAAGTAGGTT CCAATC
ASPA3 R ns attB2	Reverse	CDS cloning	GGGGACCACTTTGTACAAGAAA GCTGGGTCAGCAGCTTTGGCGA ATC
ASPA2 400 SEQ F	Forward	Primer for sequencing	CAATCTGGTGGTGATTCTG
ASPA2 800 SEQ F	Forward	Primer for sequencing	CTGGCAGTTCGACATGGGTG
ASPA2 1200 SEQ F	Forward	Primer for sequencing	TTGAGGCAGAACATGACTCA
ASPA2 1400 SEQ F	Forward	Primer for sequencing	ACAATGTATTAGCGGCTTTA
ASPA1 400 SEQ F	Forward	Primer for sequencing	AGAAGAATGGAAAAGCTGCC
ASPA1 800 SEQ F	Forward	Primer for sequencing	TGTTCTTATTGGCGGTGCAC
ASPA1 1200 SEQ F	Forward	Primer for sequencing	GAGCGCATATTGAACTACGT

ASPA1 1400 SEQ F	Forward	Primer for sequencing	TGCTCTTGACGTTGCTCCAC
ASPA3 400 SEQ F	Forward	Primer for sequencing	AGTCATCGTCATATAGAAAG
ASPA3 800 SEQ F	Forward	Primer for sequencing	GTTTGACATGGGTGATCTCC
ASPA3 1200 SEQ F	Forward	Primer for sequencing	ACACAAGAACGCATACTCGC
ASPA3 1400 SEQ F	Forward	Primer for sequencing	TTTCACGGCAATGGATATTG
ASPA2 promoter PstI F	Forward	Promoter cloning	GGGCTGCAGATCTGATGCAAAGA CGTGAC
ASPA2 promoter SalI R	Reverse	Promoter cloning	GGGGTCGACTTTGACCTACAAAA TCAAAG
ASPA1 PRO PmeI SacI F	Forward	Promoter cloning	GAGTGTTTAAACGAGCTCAGTAA GCTTGAATGTCTTG
ASPA1 PRO SalI R	Reverse	Promoter cloning	GAGTGTCGACTTTACCTATTCATT GACAAC
ASPA3 PRO SacI F	Forward	Promoter cloning	CACCGAGCTCGGAAACGTATGCT TATGGGT
ASPA3 PRO XhoI R	Reverse	Promoter cloning	GGGGCTCGAGTTTACCTGTCAT CAAAAAC
ASPA2 PRO 500 SEQ F	Forward	Primer for sequencing	CTCAAATCCTTATTTTTGGA
ASPA2 PRO 1000 SEQ F	Forward	Primer for sequencing	AAACCTTTAGCCTATTAAAT
ASPA2 PRO 1500 SEQ	Forward	Primer for sequencing	TCATGATGACACTTTTGTTC

F			
ASPA2 PRO 1900 SEQ F	Forward	Primer for sequencing	TCGAGGAACAGTTGTCTTAG
ASPA1 PRO 500 SEQ F	Forward	Primer for sequencing	CTCAATCCAACGGTTAGTAT
ASPA1 PRO 1000 SEQ F	Forward	Primer for sequencing	TTAGGTAAGAGTTTTGTTAC
ASPA1 PRO 1500 SEQ F	Forward	Primer for sequencing	TAGCAAAAGAAGTCTTTAGT
ASPA1 PRO 1800 SEQ F	Forward	Primer for sequencing	GGTATGGTTCTCTGCTTTTT
ASPA3 PRO 500 SEQ F	Forward	Primer for sequencing	GTACCTAATGCTAAACAAAC
ASPA3 PRO 1000 SEQ F	Forward	Primer for sequencing	CATCCTAGAAGATATCTTAA
ASPA3 PRO 1500 SEQ F	Forward	Primer for sequencing	TGTGAGTGTTCTTTTATACT
ASPA3 PRO 2000 SEQ F	Forward	Primer for sequencing	TCTTAGTCTAATAGTCTTCA
mCherry SpeI F	Forward	mCherry cloning	GGGGACTAGTATGGTGAGCAAG GGCGAGGA
mCherry PstI R	Reverse	mCherry cloning	GGGGTTATAATTACTTGTACAGCT CGTCCAT

ASPA2 D107A F	Forward	Site-directed mutagenesis	CTGTCATTTTGTACCGGAAGCT CTAACC
ASPA2 D107A R	Reverse	Site-directed mutagenesis	GAGCTTCCGGTAGCAAAAATGAC AGTGAAC
ASPA2 R402Q F	Forward	Site-directed mutagenesis	GATACAGAGCCAATTGCAGCAGA ACATGACT
ASPA2 R402Q R	Reverse	Site-directed mutagenesis	CTTGAGTCATGTTCTGCTGCAATT GGCTCTG
ASPA2 N404A F	Forward	Site-directed mutagenesis	GAGCCAATTGAGGCAGGCCATG ACTCAAGAG
ASPA2 N404A R	Reverse	Site-directed mutagenesis	TCCTCTTTGAGTCATGGCCTGCC TCAATTG
attB1 SAPOSIN A3 F	Forward	Cloning	GGGGACAAGTTTGTACAAAAAA GCAGGCTAAATGGGTGATCTCCA AATTGCT
attB2 SAPOSIN A3 R ns	Reverse	Cloning	GGGGACCACTTTGTACAAGAAA GCTGGGTAAGCAGCTTTGGCGA ATCCAAC
attB1 AT3G51730 CDS F	Forward	CDS cloning	GGGGACAAGTTTGTACAAAAAA GCAGGCTAAATGGGTCTTAAAGC TGGAAC

attB2 AT3G51730 CDS R ns	Reverse	CDS cloning	GGGGACCACTTTGTACAAGAAA GCTGGGTAAGAATCAGCCAACTC CGGCT
attB1 AT5G01800 CDS F	Forward	CDS cloning	GGGGACAAGTTTGTACAAAAAA GCAGGCTAAATGGGCGGTAGATT TGGAGT
attB2 AT5G01800 CDS R ns	Reverse	CDS cloning	GGGGACCACTTTGTACAAGAAA GCTGGGTACGAATCTGCCAATGA CTCCAC
attB1 AT3G51730 PROMOTER SacI F	Forward	Promoter cloning	GGGGACAAGTTTGTACAAAAAA GCAGGCTAAGAGCTCAAGAGTG ATTGAAATGGTCT
attB2 AT3G51730 PROMOTER XhoI R	Reverse	Promoter cloning	GGGGACCACTTTGTACAAGAAA GCTGGGTACTCGAGGATTCCTGA TAAAGAAAAAAG
attB1 AT5G01800 PROMOTER SacI F	Forward	Promoter cloning	GGGGACAAGTTTGTACAAAAAA GCAGGCTAAGAGCTCAAGGCAAT AACCACTCGATG
attB2 AT5G01800 PROMOTER XhoI R	Reverse	Promoter cloning	GGGGACCACTTTGTACAAGAAA GCTGGGTACTCGAGGTTTCCTCG TGAGATCTATA

AT5G01800 PROMOTER 500 SEQ F	Forward	Primer for sequencing	CTCATCAGAATTACATCTC
AT3G51730 guideRNA 1 F	Forward	Primer for guide RNA in CRISPR	ATTGAGACGTTTGCACTCTGTGT G
AT3G51730 guideRNA 1 R	Reverse	Primer for guide RNA in CRISPR	AAACCACACAGAGTGCAAACGTC T
AT5G01800 guideRNA 1 F	Forward	Primer for guide RNA in CRISPR	ATTGCCGATTCTTCTCGAACCATT
AT5G01800 guideRNA 1 R	Reverse	Primer for guide RNA in CRISPR	AAACAATGGTTCGAGAAGAATCG G
ATG8a_CACC_F	Forward	Genomic sequence cloning	CACCATGATCTTTG CTGCTTGAA
ATG8a_R	Reverse	Genomic sequence cloning	TCAAGCAACGGTAAGAGATC
M13 Forward	Forward	Primer for sequencing	GTAAAACGACGGCCAG
M13 Reverse	Reverse	Primer for sequencing	CAGGAAACAGCTATGAC
35S SEQ F	Forward	Primer for sequencing	GACGCACAATCCCACTATCCTTCG
pUBC::CFP SEQ F	Forward	Primer for sequencing	CTCGAGTGCGGGATCCTCTA

Table S02. List of Plant PSAPLIPs. Data were screened from Uniprot. Protein ID was added if no gene ID was available in the Gene ID column. Number of SapB-like domains

was auto-predicted by Uniprot. If other type of domains were also predicted, the names of domains were indicated. After alignments, some results were added with a question mark which indicates the uncertainty of SapB-like domain numbers due to mutated or missing conserved cysteines. The inferred incomplete SapB-like domain was also indicated as incomplete? In the column. The order of domain annotations was from left to right: from N to C. Signal peptide was auto-predicted by Uniprot. NA: none available.

Species	Gene ID	Number of SapB-like domains	Signal peptide prediction
<i>Chloropocon primus</i>	A3770_07P47130	2	YES
<i>Chloropocon primus</i>	A3770_02p14820	2?	YES
<i>Chloropocon primus</i>	A3770_04p29840	2	YES
<i>Chloropocon primus</i>	A3770_04p29830	2?	YES
<i>Tetradesmus obliquus</i> (<i>Acutodesmus obliquus</i>)	BQ4739_LOCUS15020	1	NA
<i>Raphidocelis subcapitata</i>	Rsub_10640	3	YES
<i>Monoraphidium neglectum</i>	MNEG_12603	2?	YES
<i>Coccomyxa subellipsoidea</i> (strain C-169)	COCSUDRAFT_45864	3	YES
<i>Chlorella variabilis</i>	CHLNCDRAFT_58828	3	YES

<i>Chlorella sorokiniana</i>	C2E21_8413	3	YES
<i>Auxenochlorella protothecoides</i> (<i>Chlorella protothecoides</i>)	APUTEX25_001631	3	YES
<i>Tetraselmis sp. GSL018</i>	TSPGSL018_26319	PPlase FKBP- type+2	YES
<i>Micromonas commoda</i> (strain RCC299 / NOUM17 / CCMP2709)	MICPUN_62224	1	YES
<i>Micromonas commoda</i> (strain RCC299 / NOUM17 / CCMP2709)	MICPUN_105899	2	YES
<i>Micromonas commoda</i> (strain RCC299 / NOUM17 / CCMP2709)	MICPUN_98458	1+disordered region	NA
<i>Ostreococcus tauri</i>	BE221DRAFT_194138	2	YES
<i>Bathycoccus prasinos</i>	Bathy10g00200	1	YES
<i>Bathycoccus prasinos</i>	Bathy07g01820	2	YES
<i>Gonium pectorale</i>	GPECTOR_69g440	3	YES
<i>Tetraena socialis</i>	TSOC_008198	3?	YES
<i>Chlamydomonas</i>	CHLRE_05g235700v5	3	YES

<i>reinhardtii</i> (<i>Chlamydomonas smithii</i>)			
<i>Chlamydomonas</i> <i>reinhardtii</i> (<i>Chlamydomonas smithii</i>)	CHLRE_02g105200v5	3	YES
<i>Chlamydomonas eustigma</i>	CEUSTIGMA_g11715.t1	3	YES
<i>Klebsormidium nitens</i> (<i>Ulothrix nitens</i>)	KFL_001110040	3	YES
<i>Chara braunii</i>	CBR_g3540	3	YES
<i>Physcomitrella patens</i> <i>subsp. patens</i>	PHYPA_022478	3	YES
<i>Physcomitrella patens</i> <i>subsp. patens</i>	PHYPA_018982	3	YES
<i>Wollemia nobilis</i>	NA	3	YES
<i>Araucaria cunninghamii</i>	NA A0A0D6R2G8_ARACU	3	YES
<i>Picea sitchensis</i>	NA A9NUE1_PICSI	3	YES
<i>Picea sitchensis</i>	NA A9P228_PICSI	2	YES
<i>Amborella trichopoda</i>	AMTR_s00007p00225690	2	YES
<i>Amborella trichopoda</i>	AMTR_s00062p00198130	2	YES
<i>Cinnamomum micranthum</i> <i>f. kanehirae</i>	CKAN_01065200	2	YES

<i>Cinnamomum micranthum</i> <i>f. kanehirae</i>	CKAN_00757300	1+incomplete ?	YES
<i>Cinnamomum micranthum</i> <i>f. kanehirae</i>	CKAN_00308200	1	NA
<i>Anthurium amnicola</i>	Psapl1_1	2	YES
<i>Anthurium amnicola</i>	Sftpb_0	2	YES
<i>Anthurium amnicola</i>	Psapl1_2	2	YES
<i>Anthurium amnicola</i>	PSAP_6	2	YES
<i>Anthurium amnicola</i>	PSAP_15	2	YES
<i>Anthurium amnicola</i>	PSAPL1_3	1	YES
<i>Anthurium amnicola</i>	mgIC_0	1	YES
<i>Zostera marina</i>	ZOSMA_381G00120	2	YES
<i>Zostera marina</i>	ZOSMA_56G01350	2	YES
<i>Apostasia shenzhenica</i>	AXF42_Ash004723	2	YES
<i>Apostasia shenzhenica</i>	AXF42_Ash015547	2	YES
<i>Dendrobium catenatum</i>	MA16_Dca011512	2	YES
<i>Dendrobium catenatum</i>	MA16_Dca015668	2	YES
<i>Dendrobium catenatum</i>	MA16_Dca010547	2	YES
<i>Dendrobium catenatum</i>	MA16_Dca020165	2	YES
<i>Dendrobium catenatum</i>	MA16_Dca009510	incomplete?+ 1	YES

<i>Ensete ventricosum</i> (<i>Musa ensete</i>)	B296_00015606	2	YES
<i>Ensete ventricosum</i> (<i>Musa ensete</i>)	B296_00023675	2	YES
<i>Ensete ventricosum</i> (<i>Musa ensete</i>)	GW17_00023743	2	YES
<i>Ensete ventricosum</i> (<i>Musa ensete</i>)	B296_00030464	2	YES
<i>Musa acuminata</i> subsp. <i>malaccensis</i> (<i>Musa malaccensis</i>)	103971073	2	YES
<i>Musa acuminata</i> subsp. <i>malaccensis</i> (<i>Musa malaccensis</i>)	NA	2	YES
<i>Musa acuminata</i> subsp. <i>malaccensis</i> (<i>Musa malaccensis</i>)	103974546	2	YES
<i>Musa acuminata</i> subsp. <i>malaccensis</i> (<i>Musa malaccensis</i>)	103970701	2	YES
<i>Musa acuminata</i> subsp. <i>malaccensis</i> (<i>Musa malaccensis</i>)	103995409	2	YES

<i>malaccensis</i> (<i>Musa malaccensis</i>)			
<i>Musa acuminata</i> subsp. <i>malaccensis</i> (<i>Musa malaccensis</i>)	103992043	2	YES
<i>Musa acuminata</i> subsp. <i>malaccensis</i> (<i>Musa malaccensis</i>)	NA MORGN3	2	YES
<i>Musa acuminata</i> subsp. <i>malaccensis</i> (<i>Musa malaccensis</i>)	NA MORELO	2	NA
<i>Musa balbisiana</i>	C4D60_Mb06t00990	2	NA
<i>Musa balbisiana</i>	C4D60_Mb08t21690	2	YES
<i>Musa balbisiana</i>	C4D60_Mb11t07550	2	YES
<i>Musa balbisiana</i>	C4D60_Mb02t19540	2	YES
<i>Musa balbisiana</i>	C4D60_Mb10t19630	2	YES
<i>Musa balbisiana</i>	C4D60_Mb10t28080	2	YES
<i>Musa balbisiana</i>	C4D60_Mb07t12400	1	YES
<i>Ananas comosus</i> (<i>Ananas ananas</i>)	ACMD2_06213	2	YES
<i>Ananas comosus</i> (<i>Ananas</i>	ACMD2_06262	2	NO

<i>ananas)</i>			
<i>Phoenix dactylifera</i>	LOC103721950	2	YES
<i>Phoenix dactylifera</i>	LOC103702109	2	YES
<i>Phoenix dactylifera</i>	LOC103718784	2	YES
<i>Phoenix dactylifera</i>	LOC103713171	2	YES
<i>Phoenix dactylifera</i>	LOC103704544	1	NA
<i>Leersia perrieri</i>	NA A0A0D9UX66	2	NA
<i>Leersia perrieri</i>	NA A0A0D9WF85	2	YES
<i>Leersia perrieri</i>	NA A0A0D9UX65	2	NA
<i>Oryza barthii</i>	NA A0A0D3HQL1	2	YES
<i>Oryza barthii</i>	NA A0A0D3EK02	2	YES
<i>Oryza barthii</i>	NA A0A0D3G678	2	YES
<i>Oryza brachyantha</i>	102703271	2	YES
<i>Oryza brachyantha</i>	NA J3M622	2	NA
<i>Oryza brachyantha</i>	102702884	2	YES
<i>Oryza glaberrima</i>	NA I1QX61	2	YES
<i>Oryza glaberrima</i>	NA I1NKJ7	2	YES
<i>Oryza glaberrima</i>	NA I1PUJ1	2	YES
<i>Oryza glumipatula</i>	NA A0A0D9Y3R0	2	NA
<i>Oryza glumipatula</i>	NA A0A0D9ZXK7	2	YES
<i>Oryza glumipatula</i>	NA A0A0E0BMU7	2	NA

<i>Oryza meridionalis</i>	NA A0A0E0DQ22	1+mutated 1?	YES
<i>Oryza meridionalis</i>	NA A0A0E0BXM2	2	YES
<i>Oryza punctata</i>	NA A0A0E0JEI8	1+mutated 1?	YES
<i>Oryza punctata</i>	NA A0A0E0L159	2	YES
<i>Oryza rufipogon</i>	NA A0A0E0RCU4	2	YES
<i>Oryza rufipogon</i>	NA A0A0E0MRU5	2	YES
<i>Oryza rufipogon</i>	NA A0A0E0PKX1	2	YES
<i>Oryza sativa subsp. indica</i>	OsI_00546	2	YES
<i>Oryza sativa subsp. indica</i>	OsI_34843	2	YES
<i>Oryza sativa subsp. indica</i>	OsI_37293	2	YES
<i>Oryza sativa subsp. indica</i>	OsI_19500	2	YES
<i>Oryza sativa subsp. japonica</i>	Os12g0112200	2	YES
<i>Oryza sativa subsp. japonica</i>	P0028E10.2	2	NA
<i>Oryza sativa subsp. japonica</i>	Os01g0166700	2	NA
<i>Oryza sativa subsp. japonica</i>	Os05g0334400	2	NA
<i>Brachypodium distachyon</i>	BRADI_4g44500v3	2	YES
<i>Brachypodium distachyon</i>	BRADI_4g25580v3	2	YES

<i>Brachypodium distachyon</i>	BRADI_2g31070v3	2	YES
<i>Brachypodium distachyon</i>	BRADI_2g04110v3	2	YES
<i>Hordeum vulgare subsp. vulgare</i>	NA A0A287NI79	2	NA
<i>Hordeum vulgare subsp. vulgare</i>	NA A0A287R2L5	2	NA
<i>Hordeum vulgare subsp. vulgare</i>	NA F2DBE9	2	YES
<i>Hordeum vulgare subsp. vulgare</i>	NA F2CQA9	2	YES
<i>Hordeum vulgare subsp. vulgare</i>	NA A0A287KA24	2	YES
<i>Aegilops tauschii subsp. strangulata</i>	NA A0A453HNI0	2	YES
<i>Aegilops tauschii subsp. strangulata</i>	F755_31720	2	YES
<i>Aegilops tauschii subsp. strangulata</i>	NA A0A453E3V5	2	YES
<i>Aegilops tauschii subsp. strangulata</i>	NA A0A453ZQF0	2	YES
<i>Triticum aestivum</i>	NA A0A3B6JIC3	2	YES

<i>Triticum aestivum</i>	NA A0A3B6KEQ2	2	YES
<i>Triticum aestivum</i>	NA A0A3B6LJX9	2	YES
<i>Triticum aestivum</i>	NA A0A3B6MMT5	2	YES
<i>Triticum aestivum</i>	NA A0A3B6IRE3	2	YES
<i>Triticum aestivum</i>	NA A0A3B6EBA4	2	YES
<i>Triticum aestivum</i>	NA A0A3B5Y5L6	2	YES
<i>Triticum aestivum</i>	NA A0A3B6FHG9	2	YES
<i>Triticum aestivum</i>	NA A0A3B5Z461	2	NA
<i>Triticum aestivum</i>	NA A0A3B6A1D1	2	NA
<i>Triticum turgidum subsp. durum</i>	TRITD_1Av1G205520	2	YES
<i>Triticum turgidum subsp. durum</i>	TRITD_4Bv1G048790	2	NA
<i>Triticum turgidum subsp. durum</i>	TRITD_4Av1G152380	2	YES
<i>Triticum turgidum subsp. durum</i>	TRITD_3Av1G029160	2	YES
<i>Triticum turgidum subsp. durum</i>	TRITD_5Av1G112490	2	YES
<i>Triticum turgidum subsp. durum</i>	TRITD_5Bv1G093140	2	YES

<i>Triticum turgidum subsp. durum</i>	TRITD_1Bv1G194670	2	NA
<i>Triticum turgidum subsp. durum</i>	TRITD_3Bv1G033240	2	YES
<i>Triticum urartu</i>	TRIUR3_03527	2	YES
<i>Triticum urartu</i>	TRIUR3_22517	2	YES
<i>Triticum urartu</i>	TRIUR3_29270	2	YES
<i>Triticum urartu</i>	TRIUR3_22718	2	YES
<i>Arundo donax (Donax arundinaceus)</i>	NA A0A0A9R7P1	2	NA
<i>Arundo donax (Donax arundinaceus)</i>	NA A0A0A9RV12	2	YES
<i>Arundo donax (Donax arundinaceus)</i>	NA A0A0A9V0R9	2	YES
<i>Arundo donax (Donax arundinaceus)</i>	NA A0A0A9V254	2	NO
<i>Arundo donax (Donax arundinaceus)</i>	NA A0A0A9QNN3	1	NA
<i>Arundo donax (Donax arundinaceus)</i>	NA A0A0A9LQW1	1	NA
<i>Eragrostis curvula</i>	EJB05_31312	1+1 mutated?	YES

<i>Eragrostis curvula</i>	EJB05_03028	2	YES
<i>Eragrostis curvula</i>	EJB05_03037	2	YES
<i>Eragrostis curvula</i>	EJB05_29979	2	YES
<i>Eragrostis curvula</i>	EJB05_34950	2	YES
<i>Eragrostis curvula</i>	EJB05_29935	1	YES
<i>Eragrostis curvula</i>	EJB05_31440	1	NA
<i>Sorghum bicolor</i>	SORBI_3008G032600	2	YES
<i>Sorghum bicolor</i>	SORBI_3003G055700	2	YES
<i>Sorghum bicolor</i>	SORBI_3009G097200	1	YES
<i>Zea mays</i>	Zm00014a_038950	2	YES
<i>Zea mays</i>	Zm00014a_044659	2	YES
<i>Zea mays</i>	ZEMMB73_Zm00001d023371	2	YES
<i>Zea mays</i>	ZEMMB73_Zm00001d042734	2	YES
<i>Zea mays</i>	ZEMMB73_Zm00001d039719	2	YES
<i>Dichanthelium oligosanthes</i>	BAE44_0015216	2	YES
<i>Dichanthelium oligosanthes</i>	BAE44_0008708	2	YES
<i>Dichanthelium oligosanthes</i>	BAE44_0009052	2?	NO
<i>Panicum hallii</i> var. <i>hallii</i>	GQ55_8G009100	2	YES

<i>Panicum hallii</i> var. <i>hallii</i>	GQ55_5G490800	2	YES
<i>Panicum hallii</i> var. <i>hallii</i>	GQ55_3G007700	2	YES
<i>Panicum hallii</i> var. <i>hallii</i>	GQ55_3G333500	2	YES
<i>Panicum miliaceum</i>	C2845_PM08G04430	2	NA
<i>Panicum miliaceum</i>	C2845_PM17G00420	2	YES
<i>Panicum miliaceum</i>	C2845_PM05G21750	2	NA
<i>Panicum miliaceum</i>	C2845_PM07G00670	2	NA
<i>Panicum miliaceum</i>	C2845_PM06G26640	2	NO
<i>Setaria italica</i>	SETIT_7G327400v2	2	YES
<i>Setaria italica</i>	SETIT_5G117400v2	2	YES
<i>Setaria italica</i>	SETIT_8G015000v2	2	YES
<i>Setaria italica</i>	SETIT_3G284400v2	2	YES
<i>Setaria viridis</i>	SEVIR_7G337200v2	2	YES
<i>Setaria viridis</i>	SEVIR_5G113900v2	2	YES
<i>Setaria viridis</i>	SEVIR_8G013800v2	2	YES
<i>Setaria viridis</i>	SEVIR_3G292600v2	2	YES
<i>Aquilegia coerulea</i>	AQUCO_00400489v1	2	YES
<i>Macleaya cordata</i>	BVC80_1837g47	1	NA
<i>Macleaya cordata</i>	BVC80_9017g10	2	YES
<i>Papaver somniferum</i>	C5167_002404	2	NA
<i>Nelumbo nucifera</i>	LOC104597199	2	YES

<i>Nelumbo nucifera</i>	LOC104602464	2	YES
<i>Spinacia oleracea</i>	SOVF_050110	2	YES
<i>Actinidia chinensis</i> var. <i>chinensis</i>	CEY00_Acc01858	2	YES
<i>Actinidia chinensis</i> var. <i>chinensis</i>	CEY00_Acc00072	2	YES
<i>Actinidia chinensis</i> var. <i>chinensis</i>	CEY00_Acc08725	2	YES
<i>Actinidia chinensis</i> var. <i>chinensis</i>	CEY00_Acc01859	2	YES
<i>Camellia sinensis</i> var. <i>sinensis</i>	TEA_000122	1	NA
<i>Davidia involucrata</i>	Din_006700	2	YES
<i>Davidia involucrata</i>	Din_026378	2	YES
<i>Nyssa sinensis</i>	F0562_017856	2	YES
<i>Nyssa sinensis</i>	F0562_015152	1?	NA
<i>Artemisia annua</i>	CTI12_AA282550	2	YES
<i>Artemisia annua</i>	CTI12_AA349490	2	NA
<i>Helianthus annuus</i>	HannXRQ_Ch10g0286291	2	no
<i>Cynara cardunculus</i> var. <i>scolymus</i>	Ccrd_003008	2	YES

<i>Lactuca sativa</i>	LSAT_9X38061	2	YES
<i>Daucus carota subsp. sativus</i>	DCAR_010960	2	YES
<i>Daucus carota subsp. sativus</i>	DCAR_018655	2?	NA
<i>Doroceras hygrometricum</i>	F511_29468	2	NO
<i>Erythranthe guttata</i> (<i>Mimulus guttatus</i>)	MIMGU_mgv1a013247mg	2	YES
<i>Genlisea aurea</i>	M569_00799	2	YES
<i>Handroanthus impetiginosus</i>	CDL12_11605	2	YES
<i>Striga asiatica</i> (<i>Buchnera asiatica</i>)	STAS_21183	2	YES
<i>Striga asiatica</i> (<i>Buchnera asiatica</i>)	STAS_33432	2	NA
<i>Coffea canephora</i>	GSCOC_T00023234001	2	NA
<i>Cuscuta australis</i>	DM860_002763	2	YES
<i>Cuscuta campestris</i>	CCAM_LOCUS31065	2	YES
<i>Cuscuta campestris</i>	CCAM_LOCUS32789	2	YES
<i>Nicotiana attenuata</i>	A4A49_38798	2	YES
<i>Nicotiana attenuata</i>	A4A49_19559	2	YES

<i>Nicotiana sylvestris</i>	LOC104216406	2	YES
<i>Nicotiana sylvestris</i>	LOC104224609	2	YES
<i>Nicotiana tabacum</i>	LOC107812754	2	YES
<i>Nicotiana tabacum</i>	LOC107816607	2	YES
<i>Nicotiana tabacum</i>	LOC107792809	2	YES
<i>Nicotiana tabacum</i>	LOC107777346	2	YES
<i>Capsicum annuum</i>	LOC107843427	2	YES
<i>Capsicum annuum</i>	LOC107851224	1	NA
<i>Capsicum baccatum</i>	CQW23_24170	2	YES
<i>Capsicum baccatum</i>	CQW23_32279	1	YES
<i>Capsicum baccatum</i>	CQW23_29496	1	YES
<i>Capsicum chinense</i>	BC332_26027	2	YES
<i>Capsicum chinense</i>	BC332_31415	1	NA
<i>Solanum chacoense</i>	NA A0A0V0I0V1	2	YES
<i>Solanum chacoense</i>	NA A0A0V0HIM7	2	YES
<i>Solanum tuberosum</i>	102602502	2	YES
<i>Solanum lycopersicum</i>	NA A0A3Q7I0I0	2	YES
<i>Vitis vinifera</i>	VIT_08s0058g01030	2	NA
<i>Vitis vinifera</i>	VITISV_040420	2	NA
<i>Vitis vinifera</i>	Psapl1_1	2	NA
<i>Vitis vinifera</i>	VITISV_040421	1	NA

		INCOMPLETE ? + 1	
<i>Vitis vinifera</i>	CK203_030312	1	NA
<i>Vitis riparia</i>	NA Q9M614	1 INCOMPLETE ? + 1	NA
<i>Arachis hypogaea</i>	Ahy_B04g070055	1	NA
<i>Arachis hypogaea</i>	Ahy_B04g071408	2	NA
<i>Arachis hypogaea</i>	Ahy_A02g006469	2	YES
<i>Arachis hypogaea</i>	Ahy_B02g061501	2	YES
<i>Arachis hypogaea</i>	Ahy_A04g018839	2	YES
<i>Lupinus angustifolius</i>	TanjilG_22489	1	YES
<i>Lupinus angustifolius</i>	TanjilG_19378	1	YES
<i>Cicer arietinum</i>	LOC101491522	2	NA
<i>Cicer arietinum</i>	LOC101508260	2	YES
<i>Medicago truncatula</i>	MTR_7g072560	2	YES
<i>Medicago truncatula</i>	MtrunA17_Chr4g0013141	2	NA
<i>Medicago truncatula</i>	MTR_4g029040	2	YES
<i>Trifolium pratense</i>	L195_g026334	2	NA
<i>Trifolium subterraneum</i>	TSUD_160400	1?	YES
<i>Trifolium subterraneum</i>	TSUD_160390	1	YES

<i>Trifolium subterraneum</i>	TSUD_266660	2	YES
<i>Lotus japonicus</i>	NA I3S9R9	2	YES
<i>Cajanus cajan</i>	KK1_035920	2	YES
<i>Cajanus cajan</i>	KK1_029931	2	YES
<i>Mucuna pruriens</i>	CR513_52785	2	NA
<i>Mucuna pruriens</i>	CR513_55238	2	NA
<i>Phaseolus vulgaris</i>	PHAVU_008G084800g	2	YES
<i>Phaseolus vulgaris</i>	PHAVU_008G0847000g	2	YES
<i>Glycine max</i>	GLYMA_18G212100	2	YES
<i>Glycine max</i>	GLYMA_19G111400	1	YES
<i>Glycine max</i>	GLYMA_09G277100	2	YES
<i>Glycine max</i>	GLYMA_18G211900	2	NA
<i>Glycine max</i>	GLYMA_01G131400	2	YES
<i>Glycine max</i>	GLYMA_09G277200	2	YES
<i>Glycine max</i>	GLYMA_04G159500	1	NA
<i>Glycine max</i>	GLYMA_19G111500	1	YES
<i>Glycine soja</i>	D0Y65_025469	2	NA
<i>Glycine soja</i>	D0Y65_001396	2	NA
<i>Glycine soja</i>	D0Y65_025468	2	YES
<i>Glycine soja</i>	D0Y65_049180	2	NA
<i>Vigna angularis</i> var.	VIGAN_04117700	2	NA

<i>angularis</i>			
<i>Vigna angularis</i> var. <i>angularis</i>	VIGAN_04117800	2	YES
<i>Vigna angularis</i> var. <i>angularis</i>	VIGAN_09109000	2	YES
<i>Vigna radiata</i> var. <i>radiata</i>	LOC106758717	2	YES
<i>Vigna radiata</i> var. <i>radiata</i>	LOC106754929	2	YES
<i>Vigna radiata</i> var. <i>radiata</i>	LOC106758948	2	YES
<i>Vigna unguiculata</i>	DEO72_LG10g3244	2	YES
<i>Vigna unguiculata</i>	DEO72_LG10g3245	2	YES
<i>Vigna unguiculata</i>	DEO72_LG8g1152	2	YES
<i>Citrus unshiu</i>	CUMW_001140	1	NA
<i>Acer yangbiense</i>	EZV62_016774	1	YES
<i>Eucalyptus grandis</i>	EUGRSUZ_K01273	2	YES
<i>Eucalyptus grandis</i>	EUGRSUZ_A00687	2	YES
<i>Punica granatum</i>	CRG98_041613	2	YES
<i>Punica granatum</i>	CRG98_041612	2	YES
<i>Punica granatum</i>	CRG98_016680	2	YES
<i>Corchorus capsularis</i>	CCACVL1_28877	2	YES
<i>Corchorus olitorius</i>	COLO4_30004	2	YES
<i>Gossypium arboreum</i>	F383_27015	2	YES

<i>Gossypium arboreum</i>	F383_21360	2	YES
<i>Gossypium australe</i>	EPI10_020460	2	NA
<i>Gossypium barbadense</i>	GOBAR_AA23056	1	YES
<i>Gossypium barbadense</i>	GOBAR_AA12144	2	YES
<i>Gossypium barbadense</i>	GOBAR_AA02853	2	YES
<i>Gossypium barbadense</i>	ES319_D10G128500v1	2	YES
<i>Gossypium barbadense</i>	ES319_A10G160500v1	2	YES
<i>Gossypium barbadense</i>	ES319_D02G005800v1	2	YES
<i>Gossypium barbadense</i>	ES319_A02G004900v1	2	YES
<i>Gossypium darwinii</i>	ES288_D10G136900v1	2	YES
<i>Gossypium darwinii</i>	ES288_A10G179500v1	2	YES
<i>Gossypium darwinii</i>	ES288_A02G005100v1	2	YES
<i>Gossypium darwinii</i>	ES288_D02G001700v1	2	YES
<i>Gossypium hirsutum</i>	LOC107896756	2	YES
<i>Gossypium hirsutum</i>	LOC107914554	2	YES
<i>Gossypium hirsutum</i>	LOC107935966	2	YES
<i>Gossypium hirsutum</i>	LOC107903579	2	YES
<i>Gossypium mustelinum</i>	E1A91_D10G132600v1	2	YES
<i>Gossypium mustelinum</i>	E1A91_A10G165200v1	2	YES
<i>Gossypium mustelinum</i>	E1A91_D02G005900v1	2	YES
<i>Gossypium mustelinum</i>	E1A91_A02G005100v1	2	YES

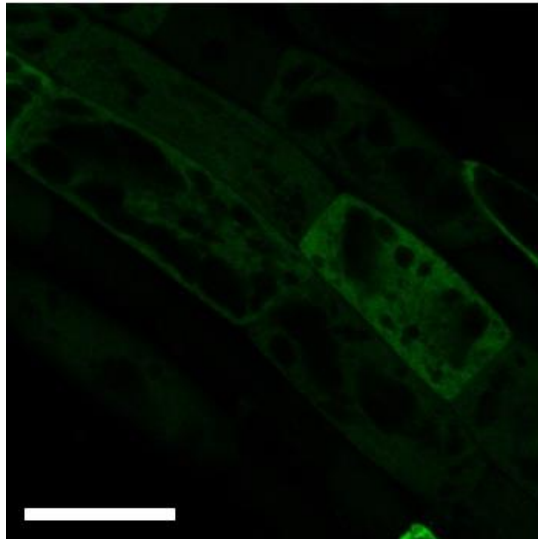
<i>Gossypium raimondii</i>	B456_011G129400	2	YES
<i>Gossypium raimondii</i>	B456_005G005800	2	YES
<i>Gossypium tomentosum</i>	ES332_D10G139500v1	2	YES
<i>Gossypium tomentosum</i>	ES332_A10G178500v1	2	YES
<i>Gossypium tomentosum</i>	ES332_A02G005200v1	2	YES
<i>Gossypium tomentosum</i>	ES332_D02G005900v1	2	YES
<i>Theobroma cacao</i>	TCM_019744	2	YES
<i>Arabis alpina</i>	AALP_AA5G141700	2	YES
<i>Arabis nemorensis</i>	ANE_LOCUS23250	2	YES
<i>Arabis nemorensis</i>	ANE_LOCUS15826	2	YES
<i>Arabis nemorensis</i>	ANE_LOCUS15790	2	YES
<i>Brassica rapa</i> subsp. <i>pekinensis</i>	NA M4D8N0	2	YES
<i>Brassica rapa</i> subsp. <i>pekinensis</i>	NA M4CRM9	2	YES
<i>Brassica napus</i>	BnaC07g32480D	2	YES
<i>Brassica napus</i>	BnaA03g57960D	2	YES
<i>Brassica napus</i>	BnaCnng40660D	2	YES
<i>Brassica oleracea</i> var. <i>oleracea</i>	NA A0A0D3DSC3	2	YES
<i>Brassica oleracea</i> var.	NA A0A0D3DE13	2	YES

<i>oleracea</i>			
<i>Brassica oleracea</i> var. <i>oleracea</i>	NA A0A0D3EIH8	2	YES
<i>Arabidopsis lyrata</i> subsp. <i>lyrata</i>	ARALYDRAFT_486888	2	YES
<i>Arabidopsis lyrata</i> subsp. <i>lyrata</i>	ARALYDRAFT_666001	2	YES
<i>Arabidopsis thaliana</i>	AT5G01800	2	YES
<i>Arabidopsis thaliana</i>	AT3G51730	2	YES
<i>Capsella rubella</i>	CARUB_v10001904mg	2	YES
<i>Capsella rubella</i>	CARUB_v10019623mg	2	YES
<i>Eutrema halophilum</i>	NA E4MWI5	2	YES
<i>Eutrema salsugineum</i>	EUTSA_v10010716mg	2	YES
<i>Eutrema salsugineum</i>	EUTSA_v10014673mg	2	YES
<i>Noccaea caerulea</i>	LE_TR12690_c0_g1_i1_g.41286	2	YES
<i>Noccaea caerulea</i>	LC_TR4311_c0_g1_i1_g.15690	2	YES
<i>Noccaea caerulea</i>	GA_TR12421_c0_g1_i1_g.39801	2	YES
<i>Noccaea caerulea</i>	MP_TR8698_c0_g1_i1_g.27351	2	YES
<i>Noccaea caerulea</i>	MP_TR15565_c0_g1_i1_g.44534	2	YES
<i>Noccaea caerulea</i>	LC_TR7688_c0_g1_i1_g.27127	2	YES
<i>Noccaea caerulea</i>	GA_TR10503_c0_g1_i1_g.34365	2	YES

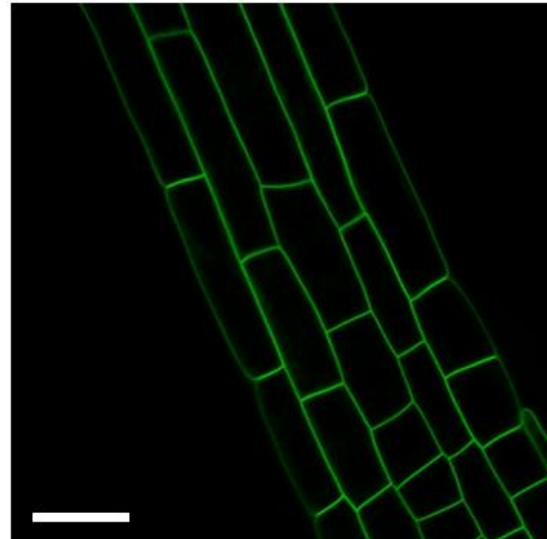
<i>Noccaea caerulea</i>	LE_TR17411_c0_g1_i1_g.56298	2	YES
<i>Rosa chinensis</i>	RchiOBHm_Chr3g0460961	2	YES
<i>Prunus persica</i>	PRUPE_6G290000	2	YES
<i>Prunus dulcis</i>	ALMOND_2B028996	2	YES
<i>Malus domestica</i>	DVH24_036312	2	NA
<i>Malus baccata</i>	C1H46_040009	2	YES
<i>Trema orientale</i>	TorRG33x02_098860	2	YES
<i>Parasponia andersonii</i>	PanWU01x14_361630	2	YES
<i>Rhizophora mucronata</i>	NA A0A2P2JI44	2	YES
<i>Populus alba</i>	D5086_0000056270	2	YES
<i>Populus trichocarpa</i>	POPTR_016G133400	2	NA
<i>Populus trichocarpa</i>	POPTR_006G107300	2	YES
<i>Juglans regia</i>	LOC108989981	2	YES
<i>Juglans regia</i>	LOC109019257	2	YES
<i>Fagus sylvatica</i>	FSB_LOCUS40270	2	NA
<i>Cucumis sativus</i>	Csa_4G331080	2	YES
<i>Cucumis melo</i> var. <i>makuwa</i>	E5676_scaffold127G001120	2	YES
<i>Cucumis melo</i> var. <i>makuwa</i>	E6C27_scaffold1166G00310	2	YES
<i>Cucumis melo</i>	LOC103502188	2	YES

Appendix F Additional Data

A

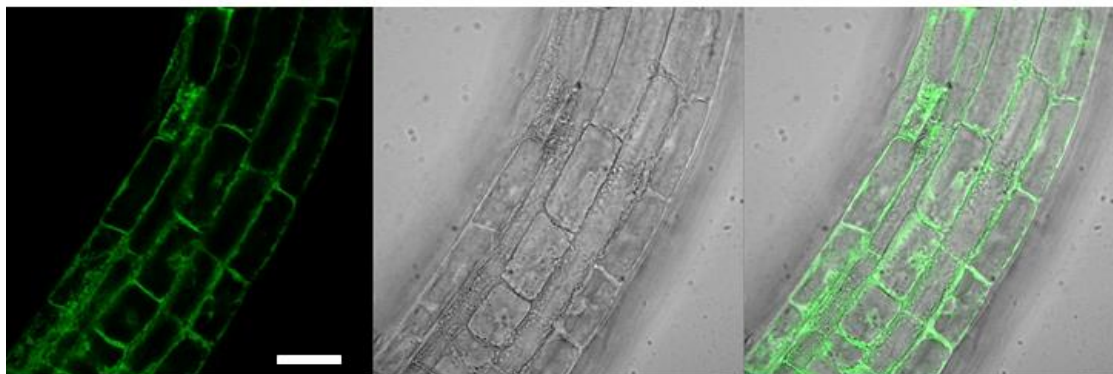


ABCB4 promoter::YFP-ABCB4



ABCB4 promoter::ABCB4-YFP

B



pUBC::ABCB4-Y1094A-YFP

Figure F-01. Subcellular localization of YFP tagged ABCB4. (A) N-terminal fusion (left) version YFP-ABCB4 and C-terminal fusion (right) version ABCB4-YFP. (B) Point mutation version ABCB4-Y1094A-YFP. Bar=20 μ m.

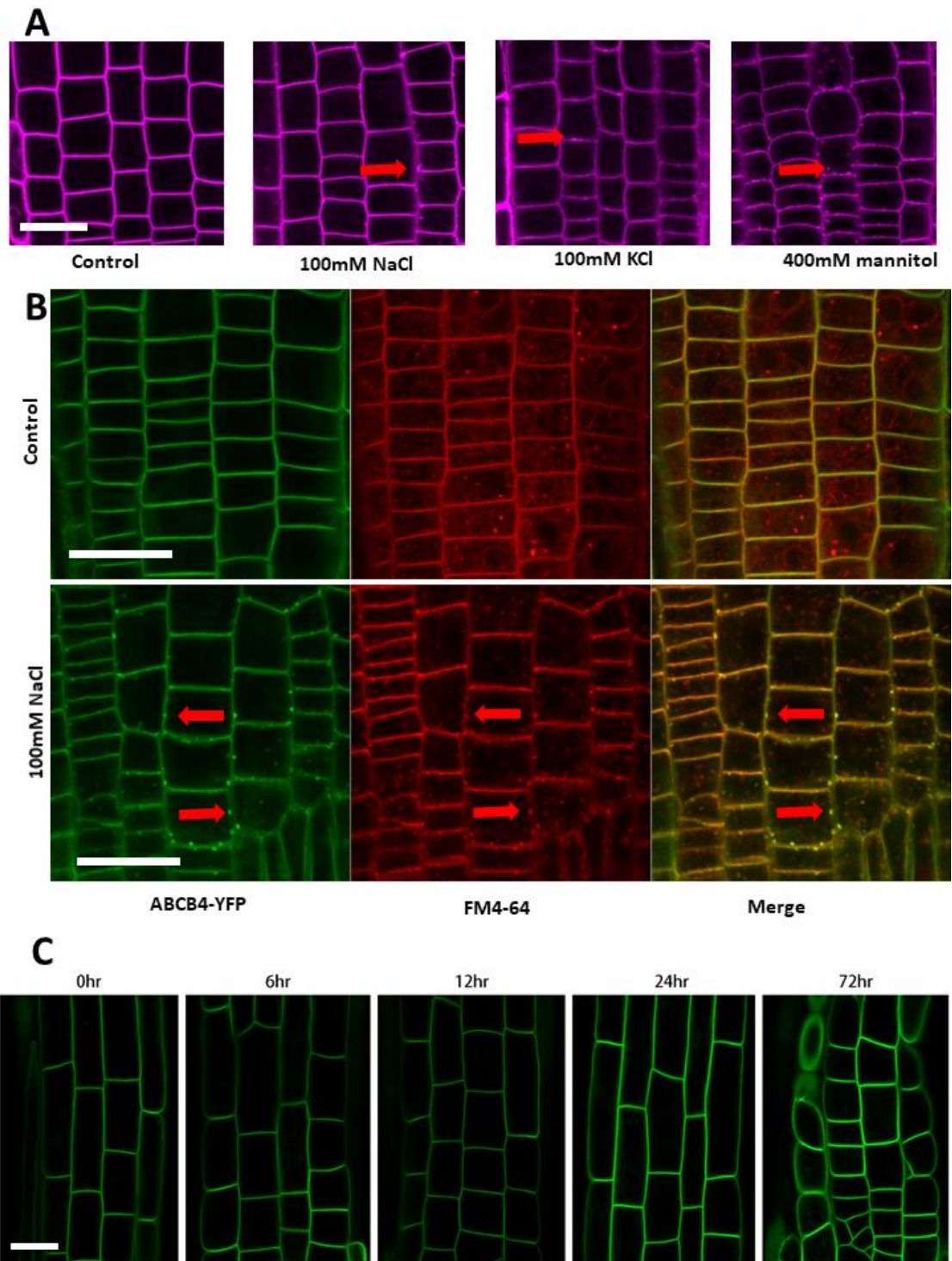


Figure F-02. ABCB4 responses to salt treatment. (A) *ABCB4* promoter::*ABCB4*-CFP responses to mock treatment, 100mM NaCl, 100mM KCl and 400mM mannitol for 20min. Arrows point to the internalized signals. Bar=20 μ m. (B) Colocalization between *ABCB4* promoter::*ABCB4*-YFP and fluorescent membrane dye FM4-64. Arrows point to the colocalized signals. Bar=20 μ m. (C) *ABCB4* promoter::*ABCB4*-YFP responses to low concentration of salt. 5DAG seedlings were treated with 25mM NaCl. Bar=20 μ m.

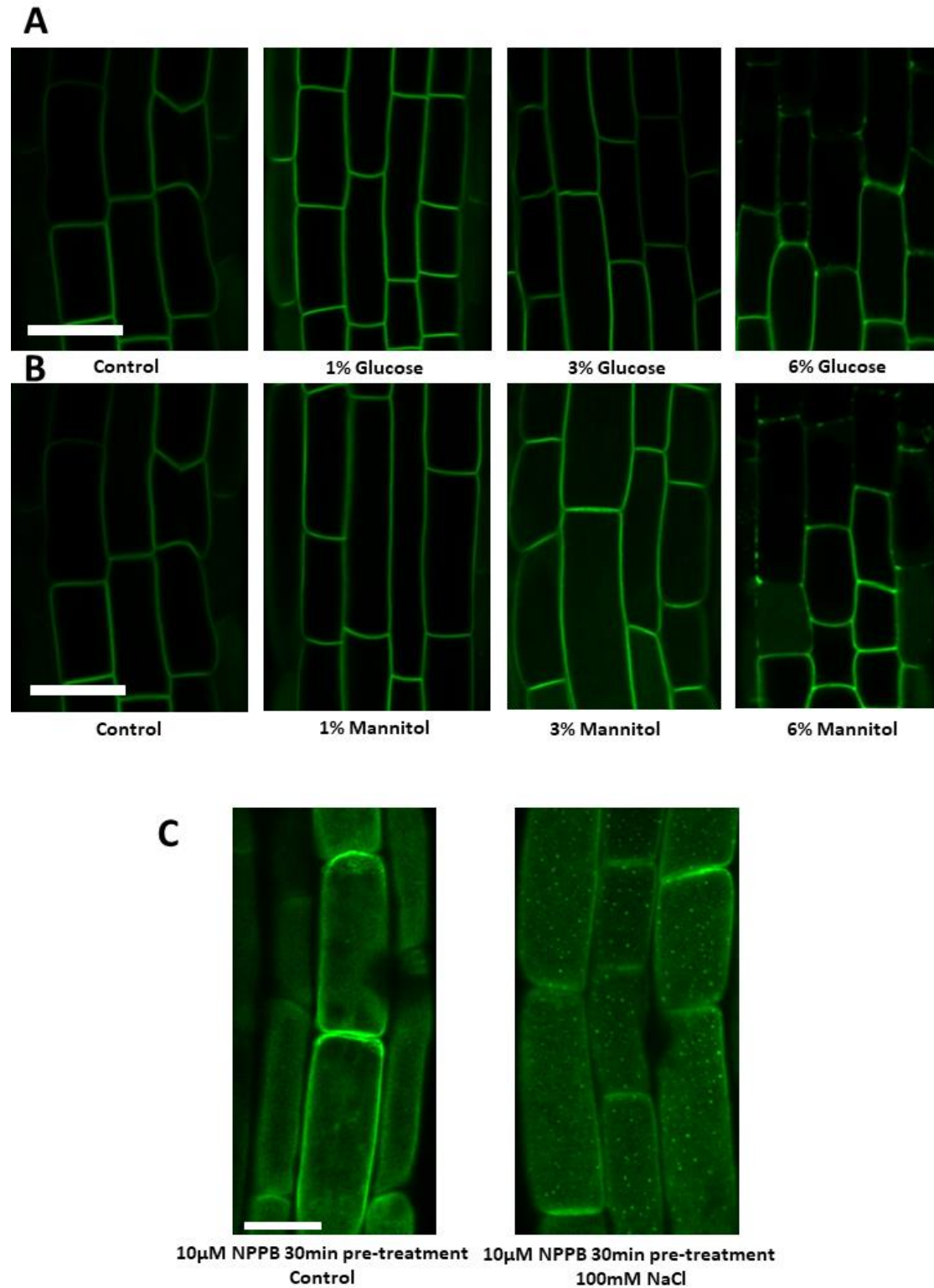


Figure F-03. Internalization of ABCB4 in response to other chemical treatments. (A) *ABCB4* promoter::*ABCB4*-YFP with glucose treatment for 20min. (B) *ABCB4* promoter::*ABCB4*-YFP with mannitol treatment for 20min. (C) Chloride inhibitor 5-nitro-2-(3-phenylpropyl-amino) benzoic acid (NPPB) pre-treated *ABCB4* promoter::*ABCB4*-YFP seedlings in response to salt treatment. Bar=20µm.

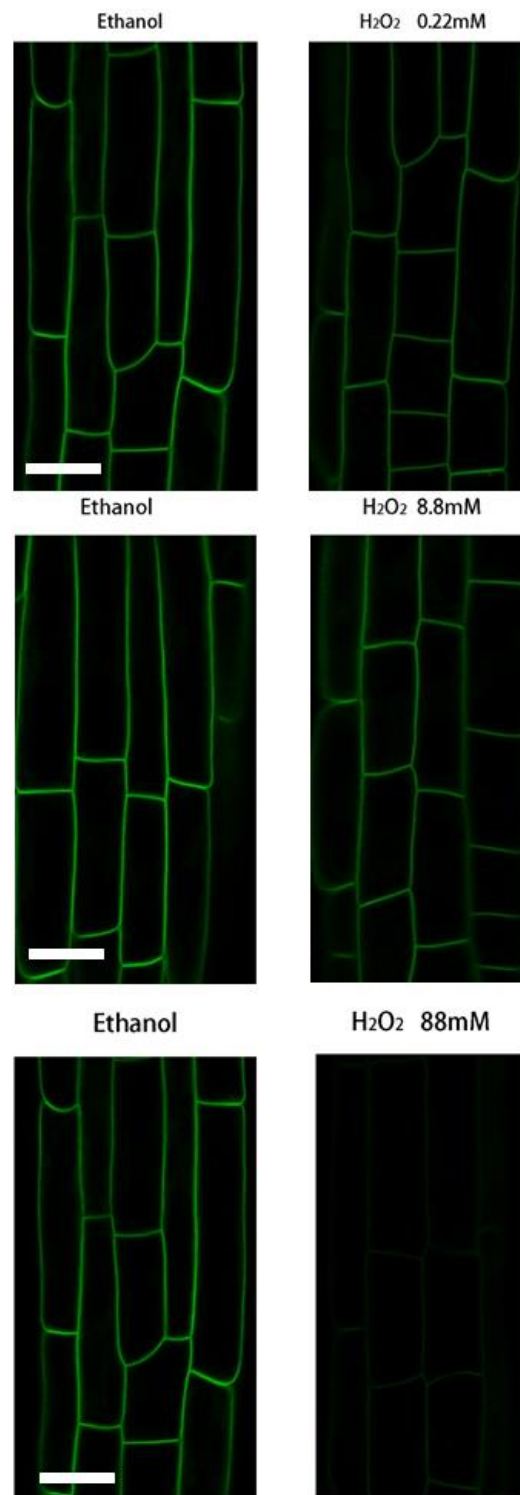


Figure F-04. *ABCB4* responses to hydrogen peroxide treatment. 5 days after germination *ABCB4* promoter::*ABCB4*-YFP seedlings were treated for 20min. Bar=20μm.

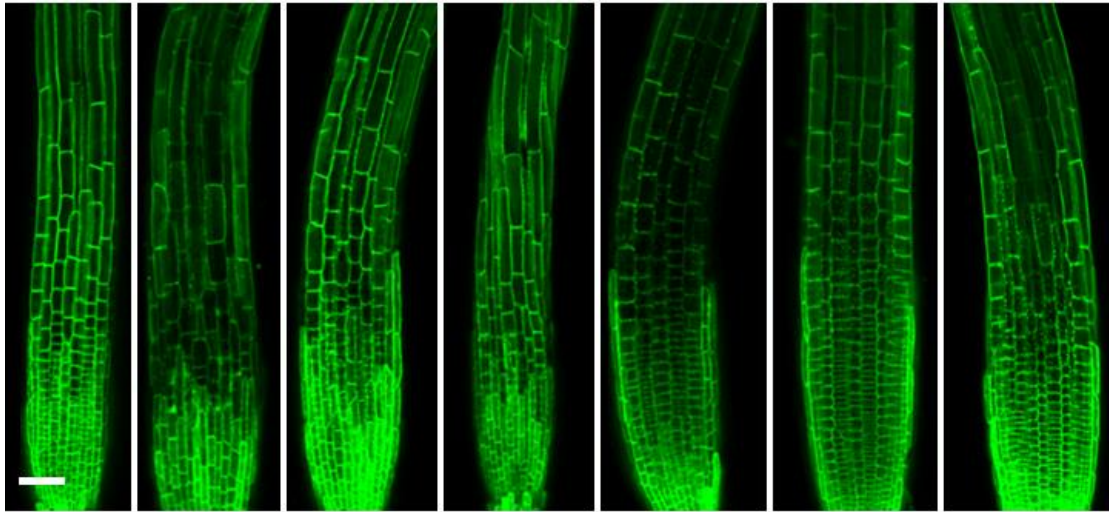
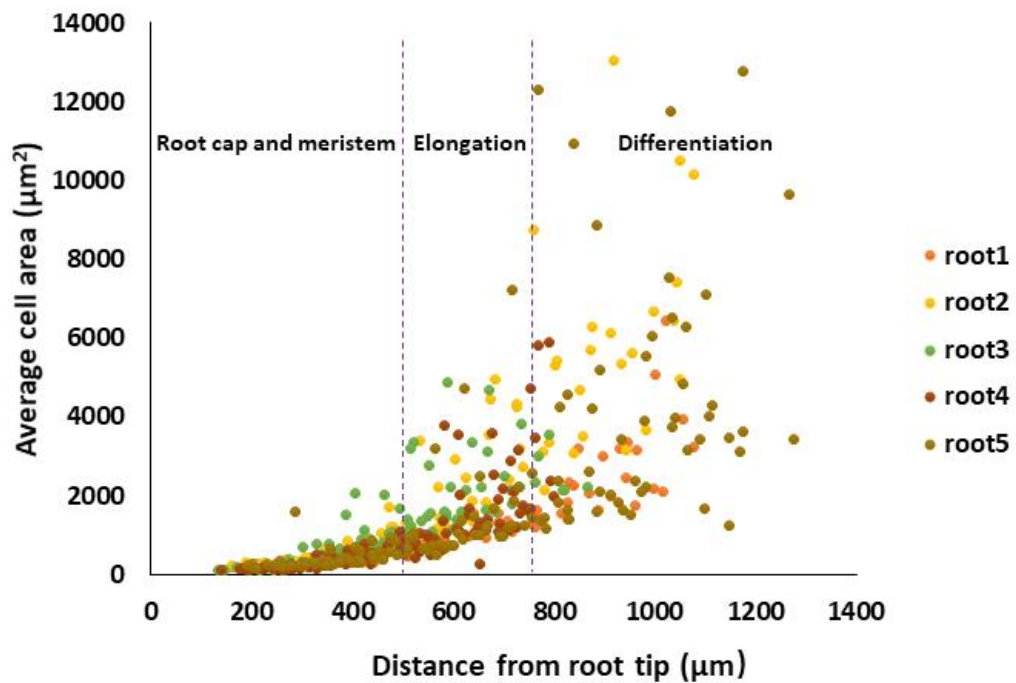
A**B**

Figure F-05. Internalization of ABCB4-YFP in response to NaCl treatment. (A) Heterogeneous responses to NaCl in different individuals. 5 days after germination *ABCB4* promoter::*ABCB4*-YFP seedlings were treated with 150mM NaCl for 20min. Bar=50 μm . (B) Cell size distribution in each section in the 5DAG seedling root tip.

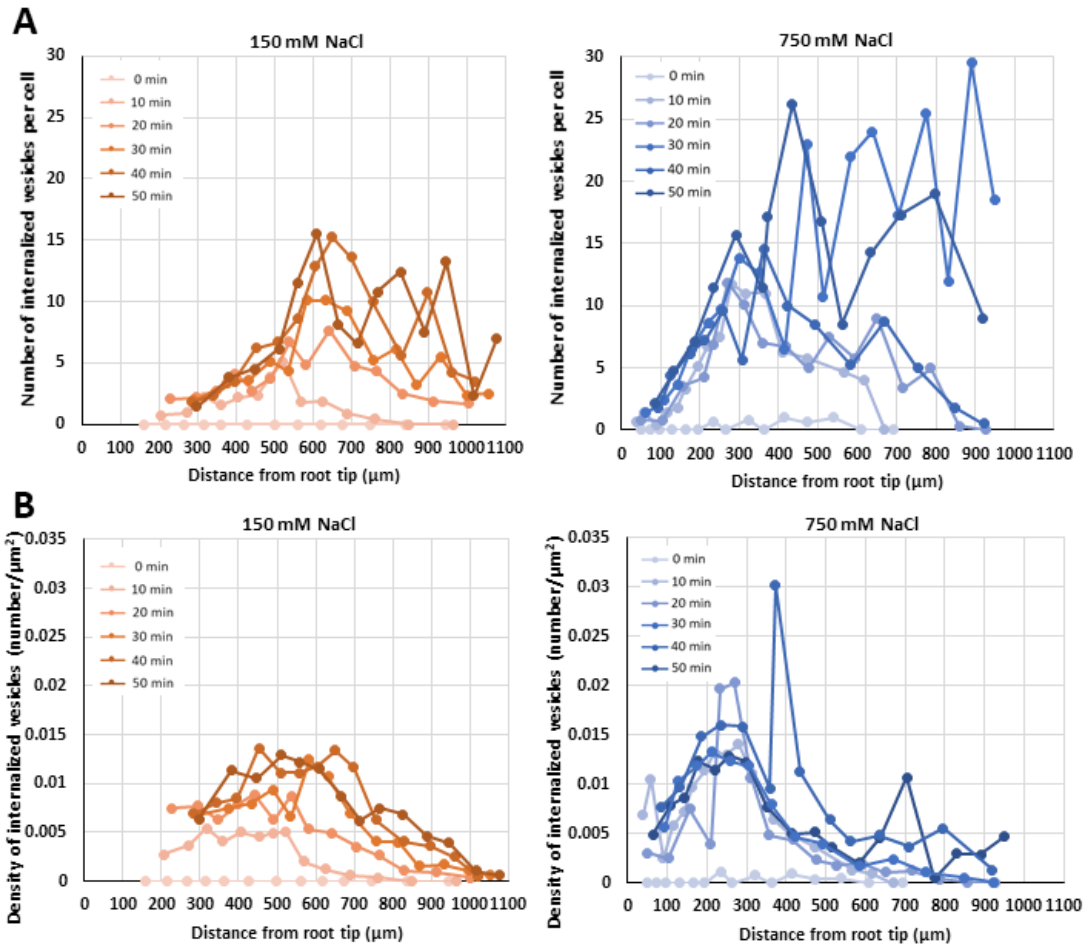


Figure F-06. ABCB4-YFP internalization events in different sections in the root tip in response to different salt concentration. (A) Numbers of internalized vesicles in each cell. (B) Density of internalized vesicles. Data from one 5 days after germination *ABCB4* promoter::*ABCB4*-YFP seedling was shown. Similar patterns also showed in other four seedlings.

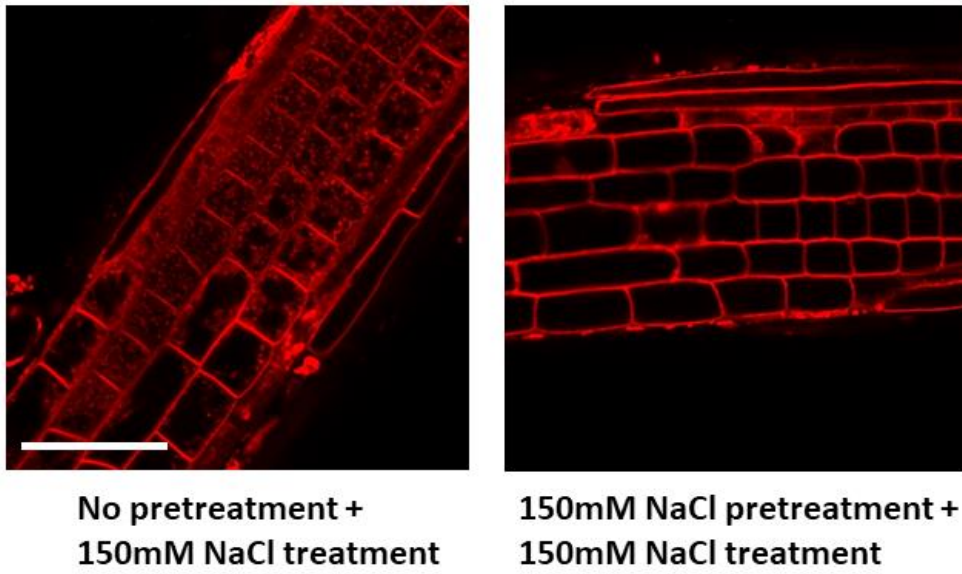


Figure F-07. NaCl pre-treated seedlings in response to salt treatment. 5 days after germination *ABCB4* promoter::*ABCB4*-YFP seedlings were treated with 150mM NaCl for 2hr (right) or no pretreatment (left), then treated with 150mM NaCl for 20min.

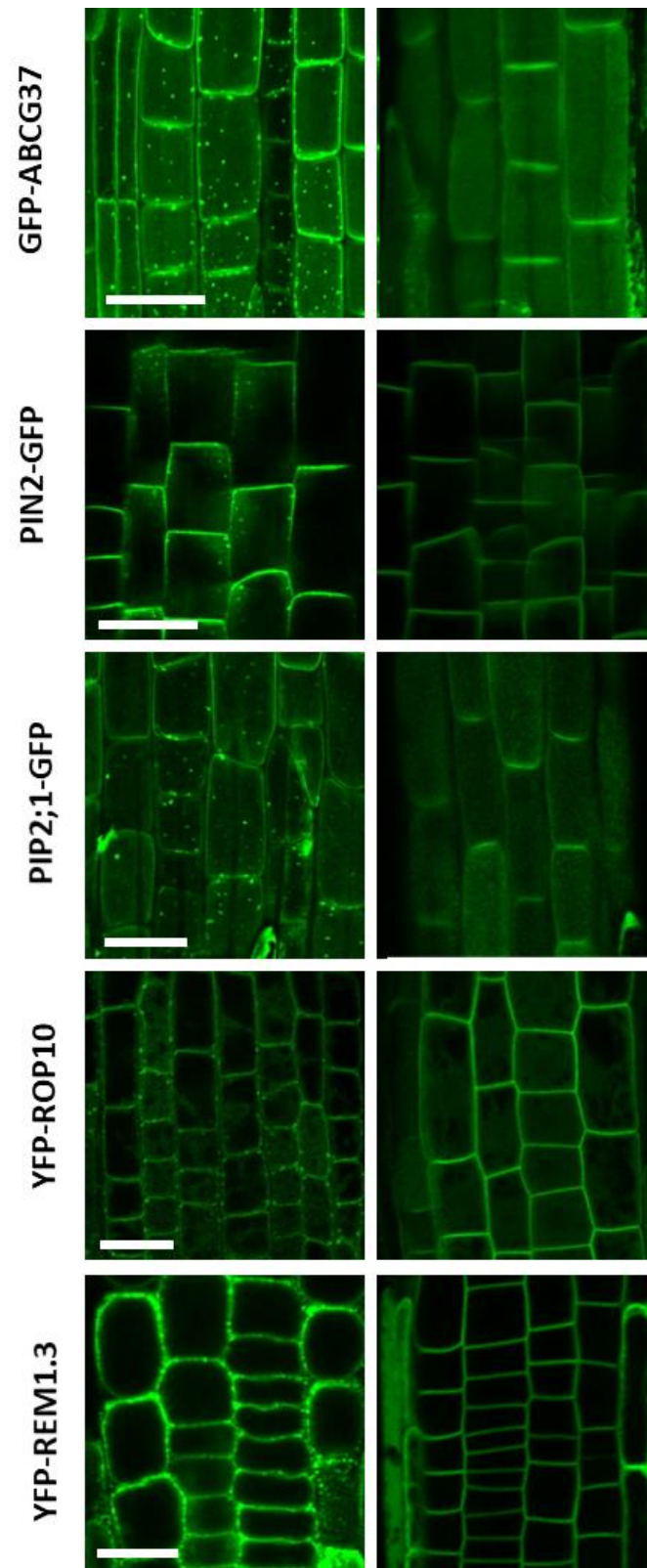


Figure F-08. Some plasma membrane markers in response to salt treatment. 5 days after germination seedlings were treated with 150mM NaCl for 20min. Bar=20μm.

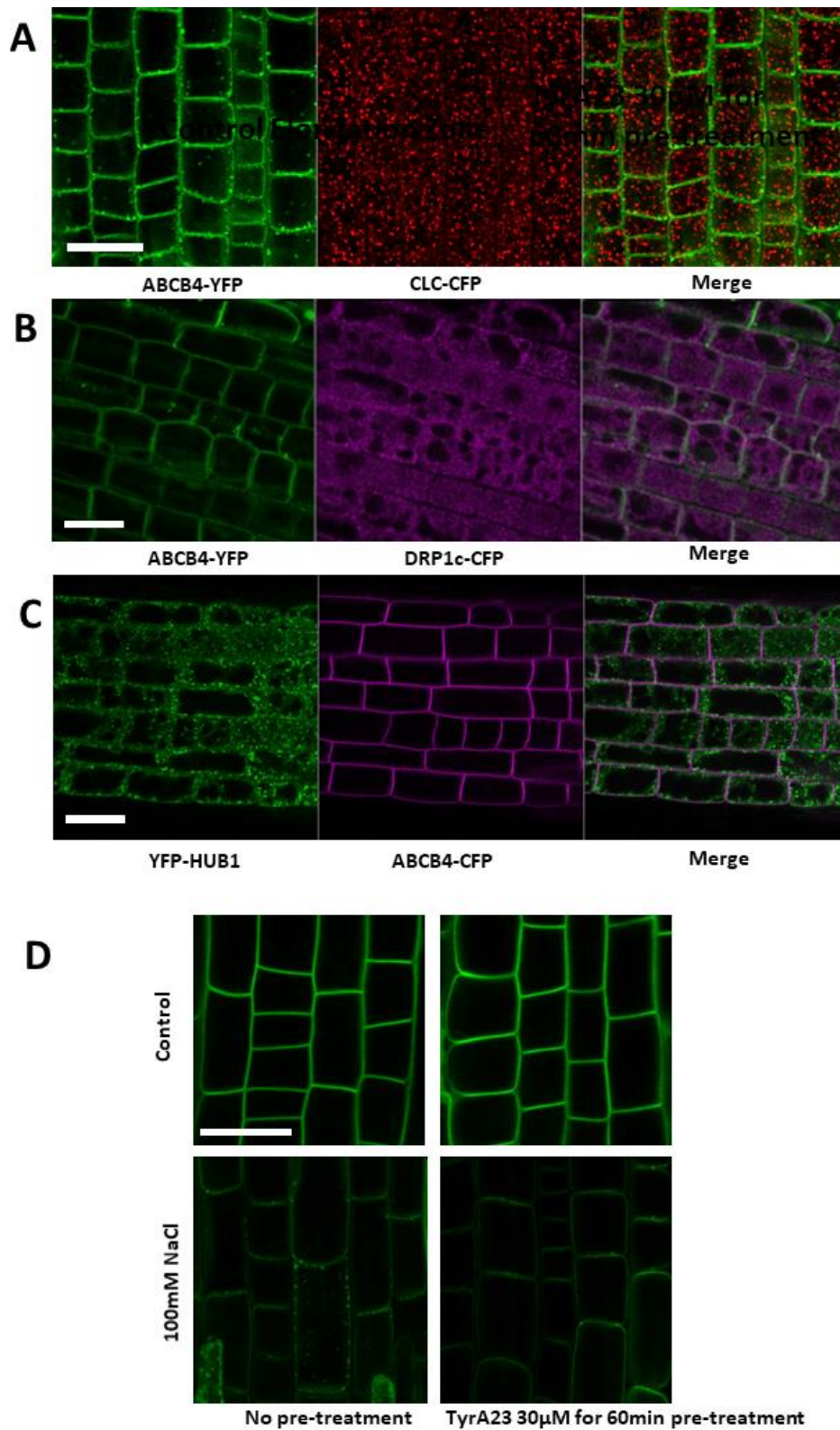


Figure F-09. Colocalization between ABCB4 and some markers associated with clathrin-mediated endocytosis. (A) Colocalization between *ABCB4* promoter::*ABCB4*-YFP with CLC-CFP. (B) Colocalization between *ABCB4* promoter::*ABCB4*-YFP with DRP1c-CFP. (C) Colocalization between *ABCB4* promoter::*ABCB4*-CFP with clathrin heavy chain HUB domain YFP-HUB1. (D) Clathrin-mediated endocytosis inhibitor

Tyrphostin A23 (TyrA23) pre-treated seedlings in response to salt treatment. 5 days after germination *ABCB4* promoter::*ABCB4*-YFP seedlings were pre-treated with TyrA23 for 60min and transferred to 100mM NaCl for 30min. Bar=20μm.

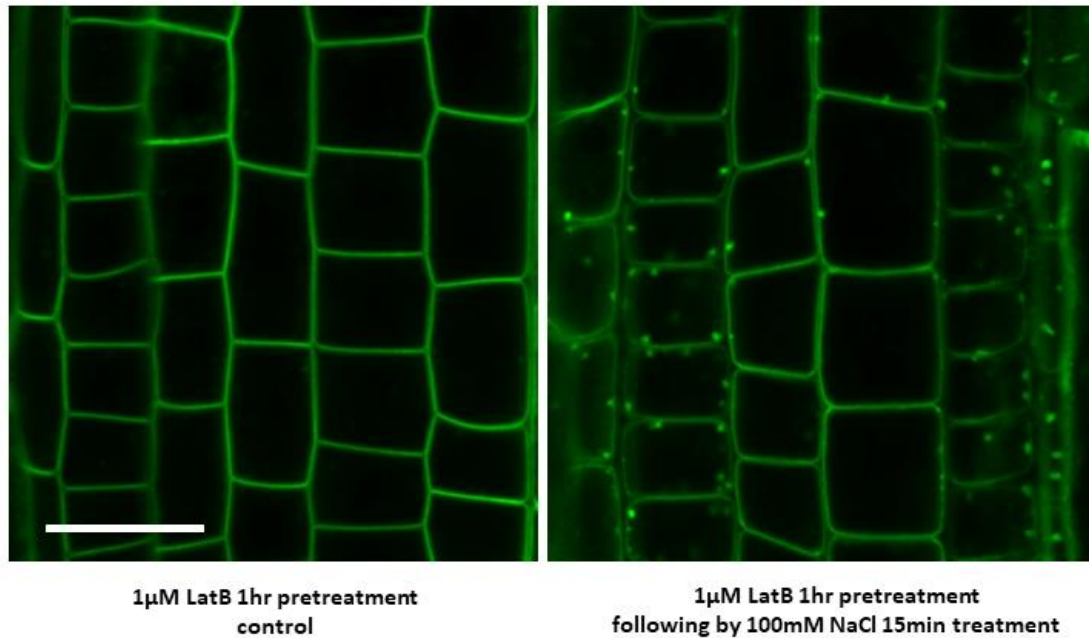


Figure F-10. *ABCB4* promoter::*ABCB4*-YFP seedlings pretreated with Latrunculin B (LatB) in response to salt treatment. 5 days after germination *ABCB4* promoter::*ABCB4*-YFP seedlings were pretreated with actin inhibitor LatB for 60min then treated with 100mM NaCl for 15min. Bar=20μm.

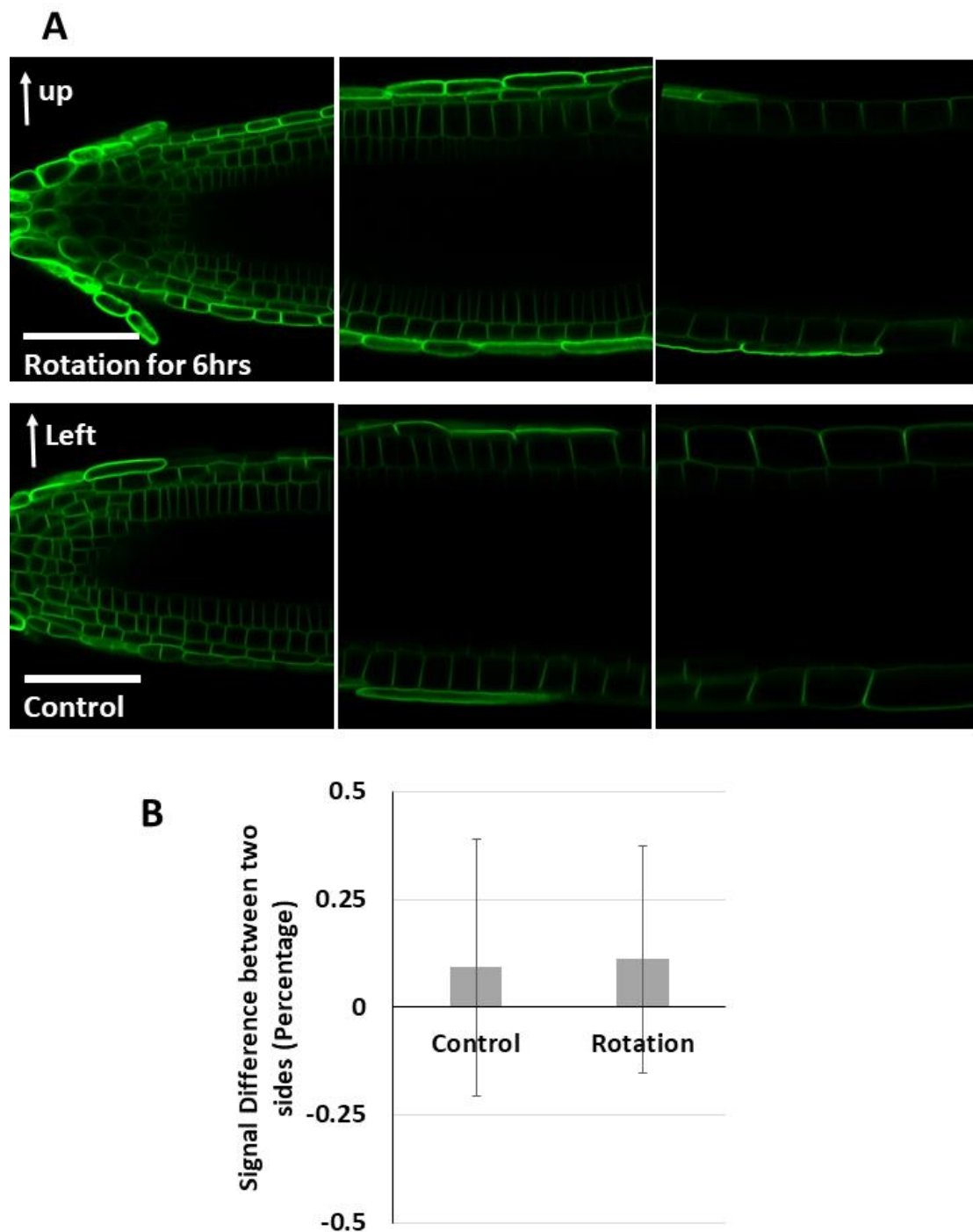


Figure F-11. ABCB4-YFP in response to gravitropism. (A) 5 days after germination vertical growing *ABCB4* promoter::*ABCB4*-YFP seedlings were rotated for 90 degree in the dark (top) or remained vertical (bottom) for 6 hours. Bar=50 μ m. (B) Fluorescent intensity differences between two sides of the root (top/bottom or left/right). N=5. $P>0.05$ by Student' *t*-test.

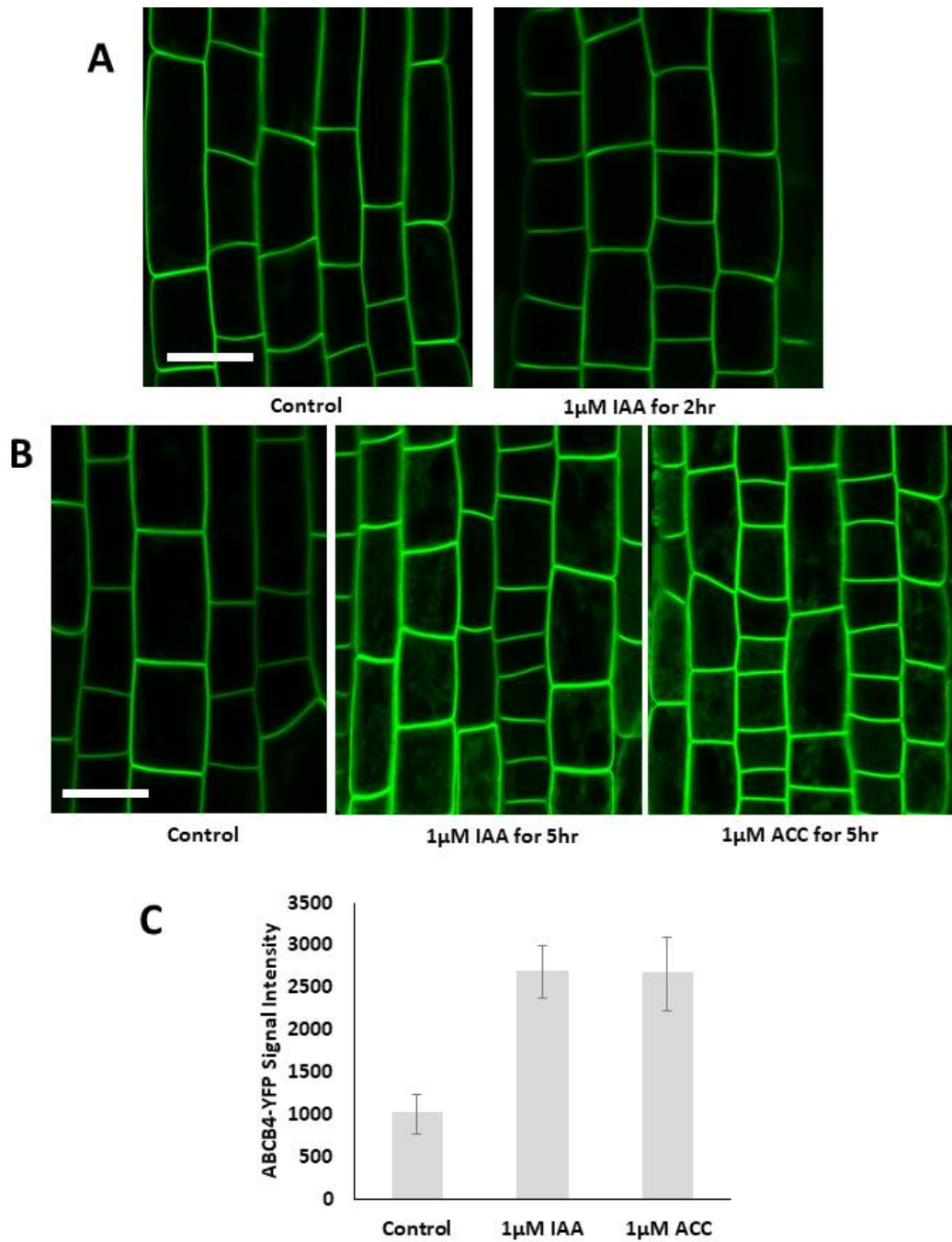


Figure F-12. Upregulation of ABCB4-YFP by IAA and ACC. (A) 5 days after germination (DAG) *ABCB4* promoter::*ABCB4*-YFP seedlings treated with IAA for 2 hr. (B) 5DAG *ABCB4* promoter::*ABCB4*-YFP seedling treated with IAA or ACC for 5 hr. (C) Fluorescent intensity statistics in (B). Bar=20μm. N=5. $P < 0.05$ by Student' *t*-test.

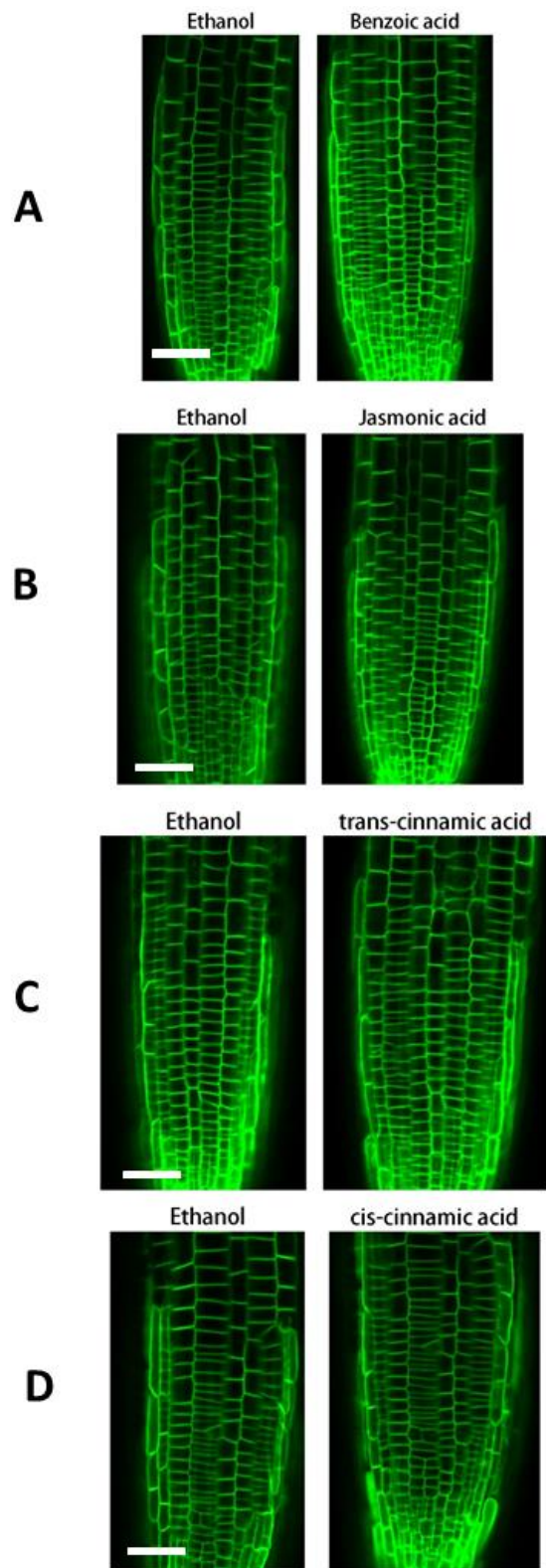


Figure F-13. ABCB4 in response to small organic acids. 5 days after germination *ABCB4* promoter::*ABCB4*-YFP seedlings were treated with the corresponding chemicals for 1 hr. Concentration = 1 μ M. Bar=50 μ m.

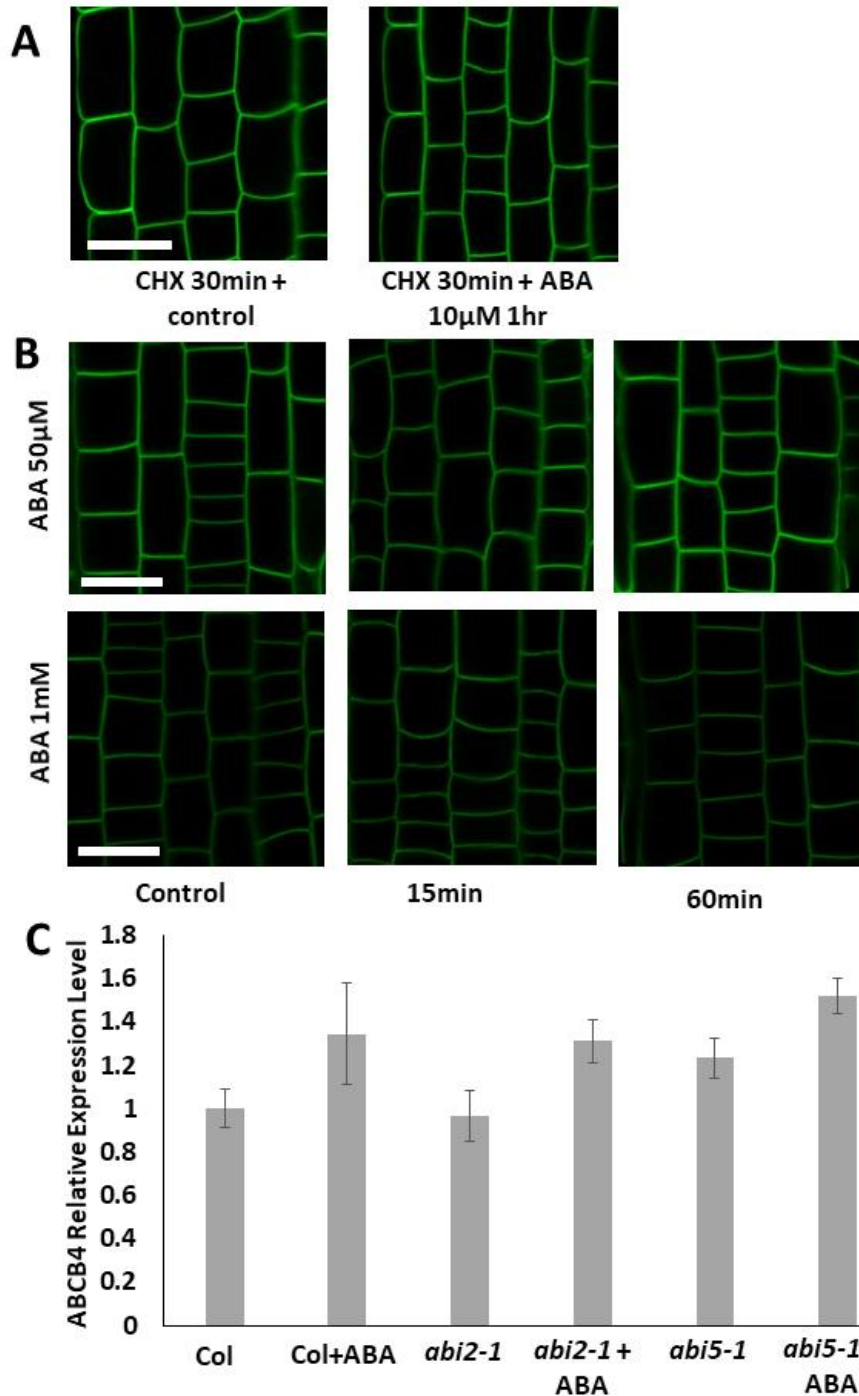


Figure F-14. ABCB4 expression in response to ABA treatment. (A) 5 days after germination (DAG) *ABCB4* promoter::*ABCB4*-YFP seedlings were pretreated with 50μM protein synthesis inhibitor cycloheximide (CHX) for 30min. Followed by 10μM ABA treatment for 1 hr. (B) ABCB4 in response to high concentration of ABA. 5DAG *ABCB4* promoter::*ABCB4*-YFP seedlings were treated with 50μM or 1mM ABA. Bar=20μm. (C) RT-PCR of ABCB4 expression in Col, *abi2-1* and *abi5-1* in response to ABA treatment. 5DAG seedlings were treated with 2μM ABA for 2 days. ACTIN2 was chosen as the reference gene. Three biological replicates.

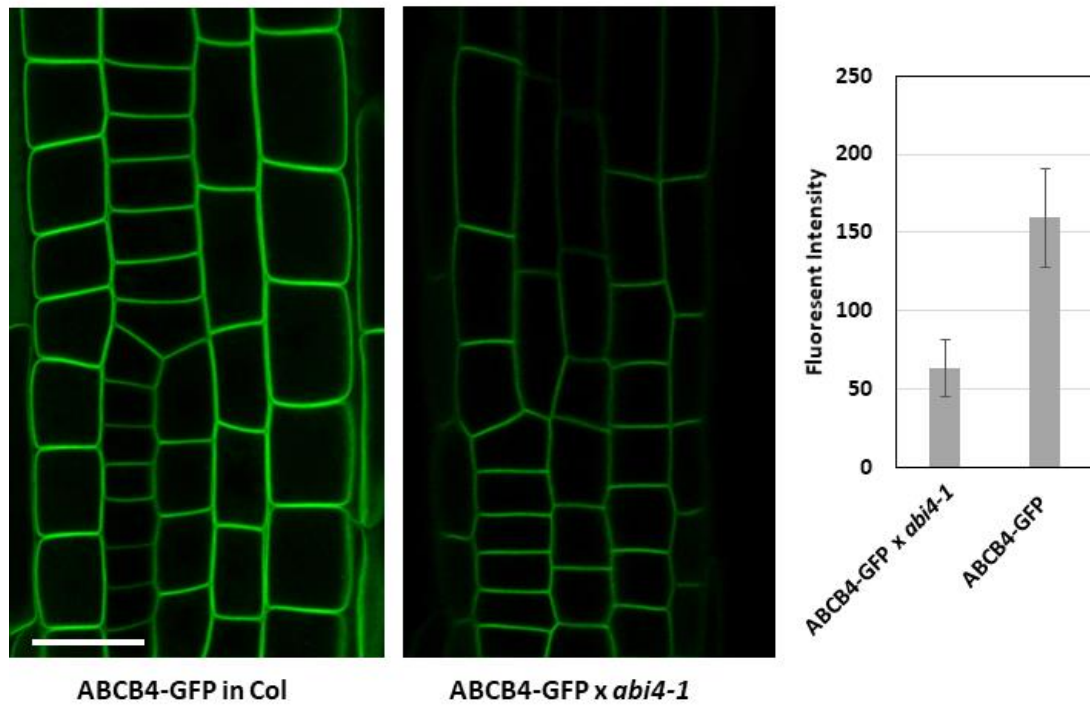


Figure F-15. *ABCB4* promoter::ABCB4-GFP in *abi4-1*. 5 days after germination seedlings were imaged. Fluorescent intensity was summarized in the right chart. Bar=20 μ m. N=5. $P < 0.05$ by Student' *t*-test.

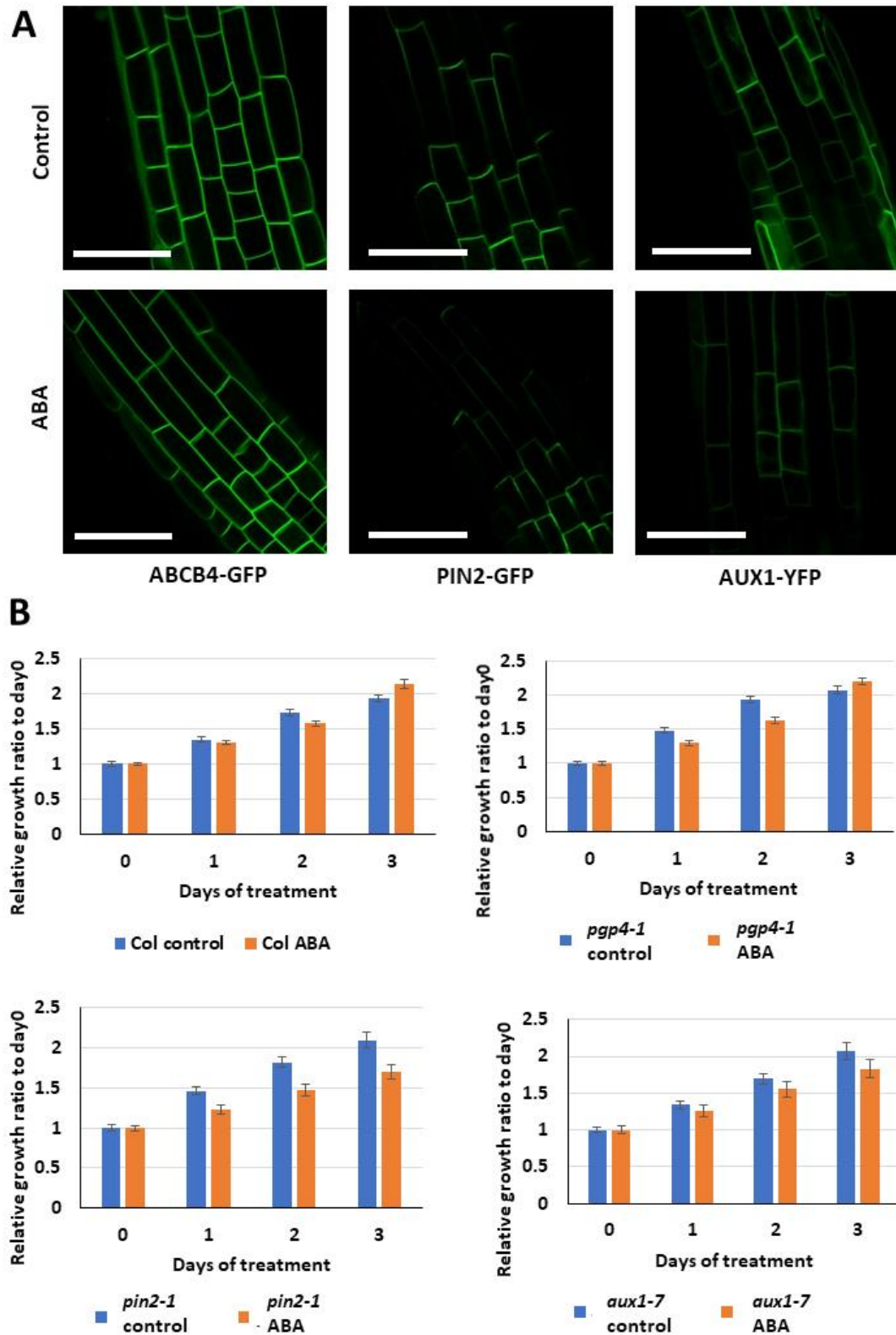


Figure F-16. ABCB4, PIN2 and AUX1 in response to ABA. (A) Confocal microscopy images of 5 days after germination seedlings treated with 1 μ M ABA for 2 hr. (B) Root growth in response to ABA treatment in *pgp4-1*, *pin2-1*, *aux1-7*. 5 days after germination seedlings were transferred to new media supplemented with 1 μ M ABA. The ratio is root length in the corresponding day / root length in day 0. N=10.

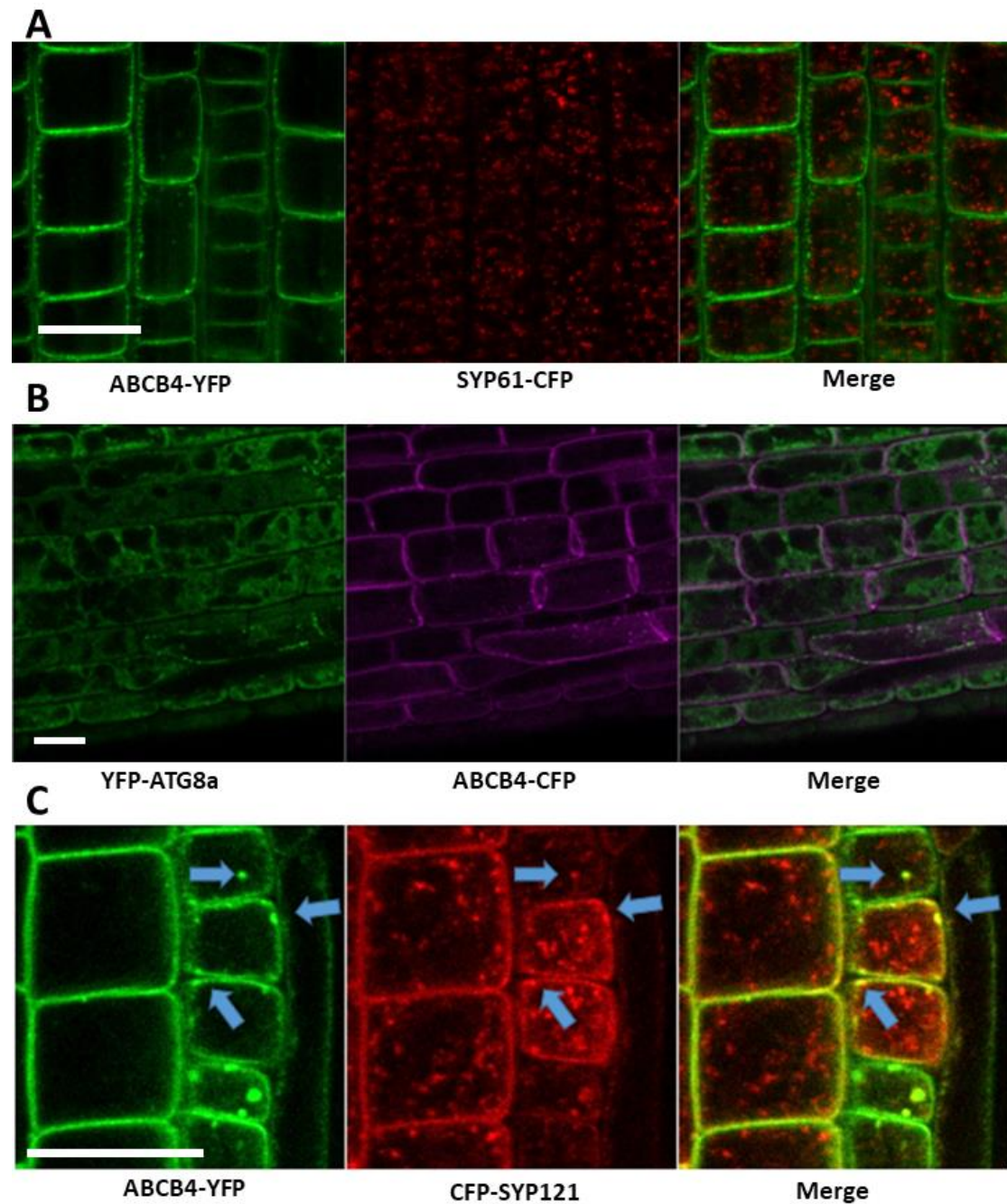


Figure F-17. Colocalization between ABCB4 and some cellular markers. 5DAG seedlings were treated with 150mM NaCl for 30min. (A) Colocalization between *ABCB4* promoter::*ABCB4*-YFP and trans-Golgi network marker SYP61-CFP. (B) colocalization between *ABCB4* promoter::*ABCB4*-CFP and autophagy marker YFP-ATG8a. (C) Colocalization between *ABCB4* promoter::*ABCB4*-YFP and plasma membrane associated SNARE protein CFP-SYP121. Arrows point to the colocalized signals. Bar=20μm.

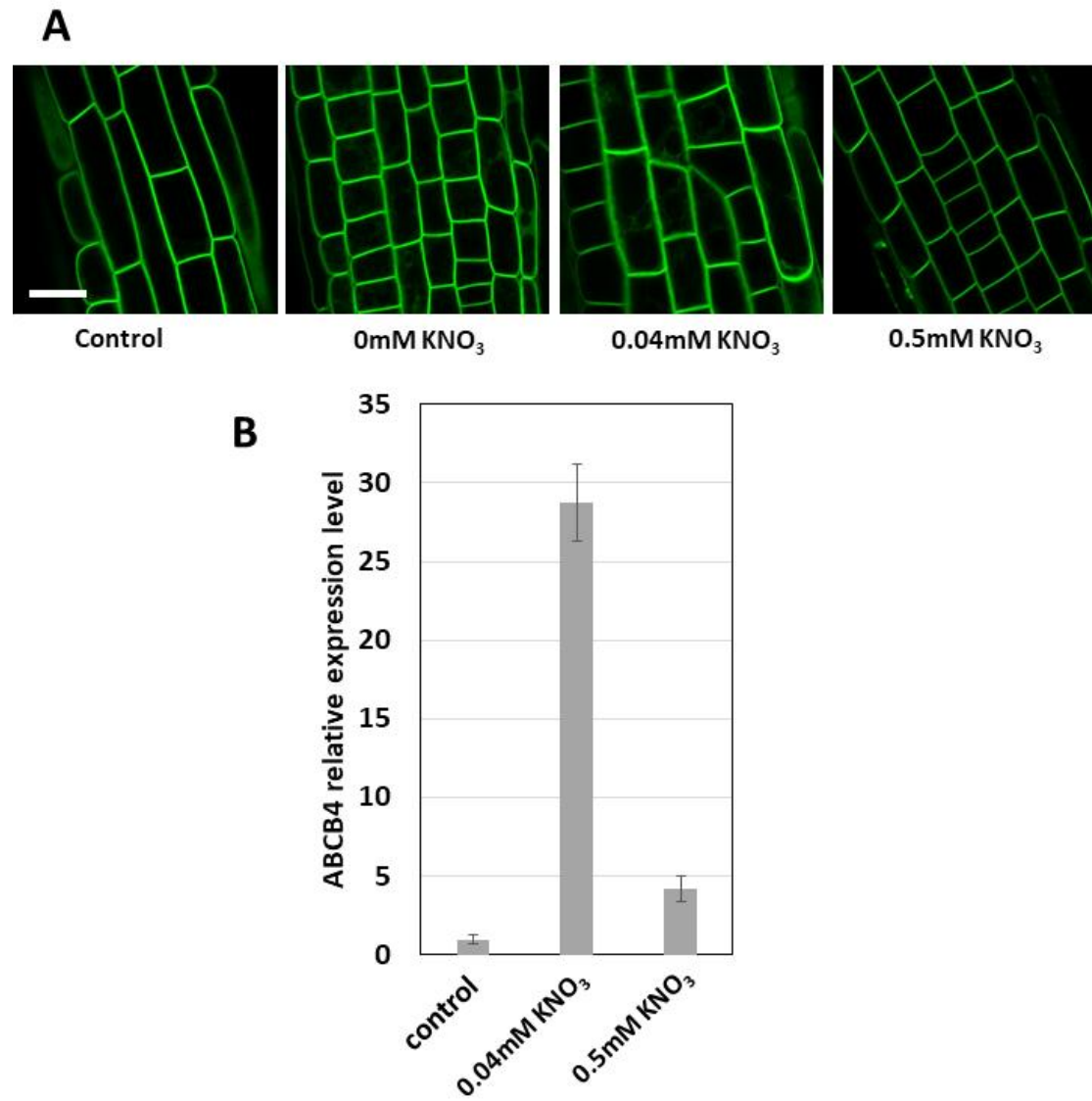


Figure F-18. ABCB4 in response to nitrogen starvation. (A) *ABCB4* promoter::*ABCB4*-YFP in response to low nitrogen treatment. *ABCB4* promoter::*ABCB4*-YFP seedlings were grown on 1/4MS media for 5 days and then transferred to new media with the corresponding nitrogen supplies for another 2 days. Bar=20 μ m. (B) RT-PCR for *ABCB4* transcript in response to low nitrogen. Wild type seedlings were grown on 1/4MS media for 5 days and then transferred to new media with the corresponding nitrogen supplies for another 2 days and harvested for RNA extraction. Three biological replicates. ACTIN2 as the reference gene.

References

- Adams, D.R., Ron, D. & Kiely, P.A. (2011). RACK1, A multifaceted scaffolding protein: Structure and function. *Cell Commun Signal*, 9: 22.
- Ahn, V.E., Faull, K.F., Whitelegge, J.P., Fluharty, A.L., Prive G.G. (2003). Crystal structure of saposin B reveals a dimeric shell for lipid binding. *Proc. Natl. Acad. Sci. U.S.A.*, 100: 38–43.
- Ahn, V.E., Leyko, P., Alattia, J.R., Chen, L., Privé, G.G. (2006). Crystal structures of saposins A and C. *Protein Sci.*, 15: 1849–1857.
- An, C., Fukusaki, E. & Kobayashi, A. (2002). Aspartic proteinases are expressed in pitchers of the carnivorous plant *Nepenthes alata* Blanco. *Planta*, 214(5): 661-667.
- Asakura, T., Watanabe, H., Abe, K. and Arai, S. (1995). Rice aspartic proteinase, oryzasin, expressed during seed ripening and germination, has a gene organization distinct from those of animal and microbial aspartic proteinases. *European Journal of Biochemistry*, 232: 77–83.
- Atta, S., Maltese, S. and Cousin, R. (2004). Protein content and dry weight of seeds from various pea genotypes. *Agronomie*, 24(5): 257–266.
- Azuma, N., O'Brien, J. S., Moser, H. W., Kishimoto, Y. (1994). Stimulation of acid ceramidase activity by saposin D. *Arch. Biochem. Biophys.*, 311: 354-357.
- Baud, S., Dubreucq, B., Miquel, M., Rochat, C. and Lepiniec, L. (2008). Storage reserve accumulation in *Arabidopsis*: metabolic and developmental control of seed filling. *The Arabidopsis book*, 6, e0113.

- Berent, S. L., Radin, N. S. (1981). Mechanism of activation of glucocerebrosidase by co- β -glucosidase (glucosidase activator protein). *Biochim. Biophys. Acta*, 664: 572-582.
- Borner, G.H., Sherrier, D.J., Weimar, T., Michaelson, L.V., Hawkins, N.D., Macaskill A., Napier, J.A., Beale, M.H., Lilley, K.S., Dupree P. (2005). Analysis of detergent-resistant membranes in *Arabidopsis*. Evidence for plasma membrane lipid rafts. *Plant Physiol*, 137:104–116.
- Bowman, J.L., Drews, G.N., Meyerowitz, E.M. (1991). Expression of the *Arabidopsis* floral homeotic gene AGAMOUS is restricted to specific cell types late in flower development. *The Plant Cell*, 3 (8):749-758.
- Bradford, K. and Bowley, J. (2003). Seeds: biology, technology and role in agriculture. In M. J. Chrispeels and D. E. Sadava, eds. *Plants, Genes, and Crop Biotechnology*. Sadbury: Jones & Bartlett Publishers, pp. 2112–2239.
- Bras, M., Queenan, B., Susin, S.A. (2005). Programmed cell death via mitochondria: different modes of dying. *Biochemistry (Moscow)*, 70:231-239.
- Brodelius, M., Hiraiwa, M., Marttila, S., Al Karadaghi, S., Picaud, S. and Brodelius, P.E. (2005). Immunolocalization of the saposin-like insert of plant aspartic proteinases exhibiting saposin C activity. Expression in young flower tissues and in barley seeds. *Physiologia Plantarum*, 125: 405-418.
- Brodersen, P., Petersen, M., Pike, H.M., Olszak, B., Skov, S., Odum, N., Jorgensen, L.B., Brown, R.E., Mundy, J. (2002). Knockout of *Arabidopsis* accelerated-cell-death11 encoding a sphingosine transfer protein causes activation of programmed cell death

and defense. *Genes Dev.*, 16 (4) :490-502.

Bruhn, H. (2005). A Short-guided tour through functional and structural features of saposin-like proteins. *Biochemical Journal*, 389: 249-257.

Bryksa, B.C., Bhaumik, P., Magracheva, E., De Moura, D.C., Kurylowicz, M., Zdanov, A., Dutcher, J.R., Wlodawer, A., Yada, R.Y. (2011). Structure and mechanism of the saposin-like domain of a plant aspartic protease. *J. Biol. Chem.*, 286: 28265-.

Bryksa, B.C., Grahame, D.A., Yada, R.Y. (2017). Comparative structure-function characterization of the saposin-like domains from potato, barley, cardoon and *Arabidopsis* aspartic proteases. *iochimica et Biophysica Acta (BBA) - Biomembranes*, 1859: 1008-1018.

Buono, A., Hudecek, R., Nowack, M. (2019). Plant proteases during developmental programmed cell death. *JOURNAL OF EXPERIMENTAL BOTANY*, 70(7): 2097–2112.

Cacas, J.L., Bure, C., Grosjean, K., Gerbeau-Pissot, P., Lherminier, J., Rombouts, Y., Maes, E., Bossard, C., Gronnier, J., Furt, F., et al. (2016). Revisiting plant plasma membrane lipids in tobacco: a focus on sphingolipids. *Plant Physiology*, 170: 367-384.

Castanheira, P., Samyn, B., Sergeant, K., Clemente, J.C., Dunn, B.M., Pires, E., Van Beeumen, J. and Faro, C. (2005). Activation, proteolytic processing, and peptide specificity of recombinant cardosin A. *The Journal of biological chemistry*, 280(13): 13047–13054.

Chen, F., Foolad, M.R. (1997). Molecular organization of a gene in barley which encodes a protein similar to aspartic protease and its specific expression in nucellar

cells during degeneration. *Plant Molecular Biology* 35: 821-831.

Chen, H., Huang, Y., Huang, G., Huang, S., Chow, T., Lin Y. (2015). Sweet potato SPAP1 is a typical aspartic protease and participates in ethephon-mediated leaf senescence. *Journal of Plant Physiology*, 180:1-17.

Chen, M., Markham, J.E., Cahoon, E.B. (2012). Sphingolipid Delta8 unsaturation is important for glucosylceramide biosynthesis and low-temperature performance in *Arabidopsis*. *The Plant Journal* 69: 769-781.

Chen, X., Pfeil, J.E. and Gal, S. (2002). The three typical aspartic proteinase genes of *Arabidopsis thaliana* are differentially expressed. *European Journal of Biochemistry*, 269: 4675-4684.

Ciaffoni, F., Salvioli, R., Tatti, M., Arancia, G., Crateri, P., Vaccaro, A.M. (2001). Saposin D solubilizes anionic phospholipid-containing membranes. *J. Biol. Chem.*, 276: 31583–31589.

Ciaffoni, F., Tatti, M., Boe, A., Salvioli, R., Fluharty, A., Sonnino, S., Vaccaro, A.M. (2006). Saposin B binds and transfers phospholipids. *J. Lipid Res.* 47: 1045–1053.

Cochrane, C.G., and Revak, S.D. (1991). Pulmonary surfactant protein B (SP-B): structure–function relationships. *Science*, 254: 566–568.

Cruz de Carvalho, M.H., d'Arcy-Lameta, A., Roy-Macauley, H., Gareil, M., El Maarouf, H., Pham-Thi, A. and Zuily-Fodil, Y. (2001). Aspartic protease in leaves of common bean (*Phaseolus vulgaris* L.) and cowpea (*Vigna unguiculata* L. Walp): enzymatic activity, gene expression and relation to drought susceptibility. *FEBS Letters*, 492.

D'Hondt, K., Bosch, D., Van Damme, J., Goethals, M., Vandekerckhove, J. and Krebbers, E. (1993). An aspartic proteinase present in seeds cleaves *Arabidopsis* 2S albumin precursors *in vitro*. *The Journal of biological chemistry*, 268(28): 20884–20891.

D'Ippólito, S., Fiol, D.F., Daleo, G.R., Guevara, M.G. (2020). Overexpression of *Arabidopsis* aspartic protease APA1 gene confers drought tolerance. *Plant Science*, 292.

Daneva, A., Gao, Z., Durme, M.V., Nowack, M.K. (2016). Functions and regulation of programmed cell death in plant development. *Annual Review of Cell and Developmental Biology*, 32: 441-468.

Daniluk, P., Lesyng, B. (2011). A novel method to compare protein structures using local descriptors. *BMC Bioinformatics* 12: 344.

Danon, A., Rotari, V.I., Gordon, A., Mailhac, N., Gallois, P. (2004). Ultraviolet-C overexposure induces programmed cell death in *Arabidopsis*, which is mediated by caspase-like activities and which can be suppressed by caspase inhibitors, P35 and Defender against Apoptotic Death. *The Journal of Biological Chemistry*, 279: 779-787.

Day, C.A., Kenworthy, A.K. (2009). Tracking microdomain dynamics in cell membranes. *Biochim. Biophys. Acta.*, 1788:245–253

De Alba, E., Weiler, S., Tjandra, N. (2003). Solution structure of human saposin C: pH-dependent interaction with phospholipid vesicles. *Biochemistry*, 42: 14729–14740.

De Moura, D. C., Bryksa, B. C., & Yada, R. Y. (2014). *In silico* insights into protein-protein interactions and folding dynamics of the saposin-like domain of *Solanum tuberosum* aspartic protease. *PloS one*, 9(9).

Devaraj, K.B., Gowda, L.R., Prakash, V. (2008). An unusual thermostable aspartic protease from the latex of *Ficus racemosa* L.. *Phytochemistry*, 69(3): 647-655.

Duarte, P., Pissarra, J., Moore, I. (2008). Processing and trafficking of a single isoform of the aspartic proteinase cardosin A on the vacuolar pathway. *Planta*, 227: 1255–1268

Egas, C., Lavoura, N., Resende, R., Brito, R. M., Pires, E., de Lima, M. C., Faro, C. (2000). The saposin-like domain of the plant aspartic proteinase precursor is a potent inducer of vesicle leakage. *J. Biol. Chem.*, 275: 38190-38196.

Egas, C., Lavoura, N., Resende, R., Brito, R.M.M., Pires, E., Pedroso de Lima, M.C., and Faro, C. (2000). The saposin-like domain of the plant aspartic proteinase precursor is a potent inducer of vesicle leakage. *J. Biol. Chem.*, 275: 38190-.

Egas, C., Lavoura, N., Resende, R., Brito, R.M., Pires, E., de Lima, M.C., Faro, C. (2000). The saposin-like domain of the plant aspartic proteinase precursor is a potent inducer of vesicle leakage. *The Journal of biological chemistry*, 275(49): 38190-38196.

Endo, M., Shimizu, H., Nohales, M. et al. (2014). Tissue-specific clocks in *Arabidopsis* show asymmetric coupling. *Nature*, 515: 419–422.

Fendrych, M., Hautegeem, T.V., Durme, M.V., Olvera-Cariilo, Y., Huysmans, M., Karimi, M., Lippens, S., J.Guerin, C., Krebs M., Schumacher, K., Nowack, M. (2014). Programmed cell death controlled by ANAC033/SOMBRERO determines root cap organ size in *Arabidopsis*. *Current Biology*, 24: 931-940.

Fendrych, M., Hautegeem, TV., Durme, MV., Olvera-Cariilo, Y., Huysmans, M., Karimi, M., Lippens, S., J.Guerin, C., Krebs, M., Schumacher, K., Nowack, M. (2014). Programmed

cell death controlled by ANAC033/SOMBRERO determines root cap organ size in *Arabidopsis*. *Current Biology*, 24: 931-940.

Fierens, K., Brijs, K., Courtin, C.M., Gebruers, K., Goesaert, H., Raedschelders, G., Robben, J., Van Campenhout., Volckaert, G. and Delcour, J.A. (2003). Molecular identification of wheat endoxylanase inhibitor TAXI-I1, member of a new class of plant proteins. *FEBS Letters*, 540

Figueiredo, R., Duarte, P., Pereira S., Pissarra J. (2006). The embryo sac of *Cynara cardunculus*: Ultrastructure of the development and localisation of the aspartic proteinase cardosin B. *Sexual Plant Reproduction*. 19: 93-101.

Frazão, C., Bento, I., Costa, C., Soares, C.M., Verissimo, P., Faro, C., Pires, E., Cooper J., Carrondo, M.A. (1999). Crystal structure of cardosin A, a glycosylated and Arg-Gly-Asp-containing aspartic proteinase from the flowers of *Cynara cardunculus* L.. *J. Biol. Chem.*, 274: 27694–27701.

Frey, M.E., D'Ippolito, S., Pepe, A., Daleo, G.R., Guevara, M.G. (2018). Transgenic expression of plant-specific insert of potato aspartic proteases (StAP-PSI) confers enhanced resistance to *Botrytis cinerea* in *Arabidopsis thaliana*. *Phytochemistry*, 149: 1-11.

Furt, F., Konig, S., Bessoule, J.J., Sargueil, F., Zallot, R., Stanislas, T., Noirot, E., Lherminier, J., Simon-Plas, F., Heilmann, I., et al. (2010). Polyphosphoinositides are enriched in plant membrane rafts and form microdomains in the plasma membrane. *Plant Physiol*, 152:2173–2187

Futerman, A.H., Hannun, Y.A. (2004). The complex life of simple sphingolipids. *EMBO Reports*, 5: 777–782.

Gao, C., Zhuang, X., Cui, Y., et al. 2015. Dual roles of an Arabidopsis ESCRT component FREE1 in regulating vacuolar protein transport and autophagic degradation. *Proc Natl Acad Sci U S A*, 112(6):1886 - 1891.

Gao, Z., Daneva, A., Salanenska, Y. et al. (2018). KIRA1 and ORESARA1 terminate flower receptivity by promoting cell death in the stigma of *Arabidopsis*. *Nature Plants*, 4: 365–375.

Gepstein, S., Sabehi, G., Carp, M.-J., Hajouj, T., Nesher, M.F.O., Yariv, I., Dor, C. and Bassani, M. (2003). Large-scale identification of leaf senescence-associated genes. *The Plant Journal*, 36: 629-642.

Glathe, S., Kervinen, J., Nimtz, M., Li, G.H., Tobin, G.J., Copeland, T.D., Ashford, D.A., Wlodawer, A. and Costa, J. (1998). Transport and activation of the vacuolar aspartic proteinase phytepsin in Barley (*Hordeum vulgare* L.). *Journal of biological chemistry*, 273(47): 31230–31236.

Gottschalk, W. and Muller, H.P. (2012). Seed proteins: biochemistry, genetics, nutritive value. *W. Gottschalk and H. P. Muller, eds., Springer Science & Business Media*.

Gruis, D., Schulze, J. and Jung, R. (2004). Storage protein accumulation in the absence of the vacuolar processing enzyme family of cysteine proteases. *The Plant cell*, 16(1): 270–290.

Gruis, D., Selinger, D.A., Curran, J.M. and Jung, R. (2002). Redundant proteolytic

mechanisms process seed 184 storage proteins in the absence of seed-type members of the vacuolar processing enzyme family of cysteine proteases. *The Plant cell*, 14(11): 2863–2882.

Guo, J., Wang, S., Valerius, O., Hall, H., Zeng, Q., Li, J.-F., Weston D.J., Brian, E. Ellis, B.E, Chen, J.-G. (2011). Involvement of *Arabidopsis* RACK1 in Protein Translation and Its Regulation by Absciscic Acid. *Plant Physiology Jan*, 155 (1): 370-383.

Hara-Nishimura, I. and Hatsugai, N. (2011). The role of vacuole in plant cell death. *Cell death and differentiation*, 18(8): 1298–304.

Hawkins, C.A., De Alba, E., Tjandra, N. (2005). Solution structure of human saposin C in a detergent environment. *J. Mol. Biol.*, 346: 1381–1392.

Hecht, O., Van Nuland, N. A., Schleinkofe,r K., Dingley, A. J., Bruhn, H., Leippe, M. and Grotzinger, J. (2004). Solution structure of the pore forming protein of *Entamoeba histolytica*. *J. Biol. Chem.*, 279: 17834–17841

Hill, C.H., Read, R.J., Deane, J.E. (2015). Structure of human saposin A at lysosomal pH. *Acta Crystallogr. F*, 71: 895–900.

Hinz, G., Hillmer, S., Baumer, M. and Hohl, I. (1999). Vacuolar storage proteins and the putative vacuolar sorting receptor BP-80 exit the Golgi apparatus of developing pea cotyledons in different transport vesicles. *The Plant cell*, 11(8): 1509–1524.

Hiraiwa, M., O'Brien, J., Kishimoto, Y., Galdzicka, M., Fluharty, A.L., Ginns, E.I., Martin, B.M. (1993). Isolation, characterization, and proteolysis of human prosaposin, the precursor of saposins (Sphingolipid Activator Proteins). *Archives of Biochemistry and*

Biophysics, 304 (1): 110-116.

Hong, Y.H., Lillehoj, H.S., Siragusa, G.R., Bannerman, D.D., Lillehoj, E.P. (2008). Antimicrobial activity of chicken NK-lysin against *Eimeria sporozoite*. *Avian Disease*, 52: 302305.

Huang, J, Zhang, T, Linstroth, L, Tillman, Z, Otegui, M.S., Owen, H.A., Zhao, D. (2016). Control of anther cell differentiation by the small protein ligand TPD1 and its receptor EMS1 in *Arabidopsis*. *PLoS Genet.* 12(8):e1006147.

Huby, E., Napier, J.A., Baillieul, F., Michaelson, L.V. and Dhondt-Cordelier, S. (2020). Sphingolipids: towards an integrated view of metabolism during the plant stress response. *New Phytol*, 225: 659-670.

Huysmans, M., Buono, R.A., Skorzinski, N., Radio, M.C., Winter, F.D., Parizot, B., Mertens, J., Karimi, M., Fendrych, M., Nowack, M.K. (2018). NAC transcription factors ANAC087 and ANAC046 control distinct aspects of programmed cell death in the *Arabidopsis* columella and lateral root cap. *The Plant Cell*, 30(9): 2197-2213.

Huysmans, M., Lema, S.A., Coll, N.S., Nowack, M.K. (2017). Dying two deaths - programmed cell death regulation in development and disease. *Current Opinion in Plant Biology*, 35: 37-44.

Huysmans, M., Buono, R.A., Skorzinski, N., Radio, M.C., Winter F.D., Parizot, B., Mertens, J., Karimi, M., Fendrych, M., Nowack, M.K. (2018). NAC transcription factors ANAC087 and ANAC046 control distinct aspects of programmed cell death in the *Arabidopsis* columella and lateral root cap. *The Plant Cell*, 30(9): 2197-2213.

- Imai, H., Ohnishi, M., Kinoshita, M., Kojima, M., Ito, S. (1995). Structure and distribution of cerebroside containing unsaturated hydroxy fatty acids in plant leaves. *Bioscience, Biotechnology, and Biochemistry*, 59: 1309–1313.
- Ines, C., Parra-Lobato, M.C., Paredes, M.A., Labrador, J., Gallardo, M., Saucedo-Garcia, M., Gavilanes-Ruiz, M., Gomez-Jimenez, M.C. (2018). Sphingolipid distribution, content and gene expression during olive-fruit development and ripening. *Frontiers in Plant Science*, 9: 28.
- Jiang, L., Phillips, T.E., Hamm, C.A., Drozdowicz, Y.M., Rea, P.A., Maeshima, M., Rogers, S.W. and Rogers, J.C. (2001). The protein storage vacuole: a unique compound organelle. *The Journal of cell biology*, 155(6): 991–1002.
- John, M., Wendeler, M., Heller, M., Sandhoff, K., Kessler, H. (2006). Characterization of human saposins by NMR spectroscopy. *Biochemistry*, 45: 5206–5216.
- Kang, S. J. and Cresswell, P. (2004). Saposins facilitate CD1d-restricted presentation of an exogenous lipid antigen to T cells. *Nat. Immunol.* 5: 175–181.
- Keinath, N.F., Kierszniowska, S., Lorek, J., Bourdais, G., Kessler, S.A., Shimosato-Asano, H., Grossniklaus, U., Schulze, W.X., Robatzek, S., Panstruga, R. (2010). PAMP (pathogen-associated molecular pattern)-induced changes in plasma membrane compartmentalization reveal novel components of plant immunity. *J. Biol. Chem.*, 285:39140–39149.
- Kerr, J.F., Wyllie, A.H., Currie, A.R. (1972). Apoptosis: a basic biological phenomenon with wide-ranging implications in tissue kinetics. *British Journal of Cancer*, 26: 239-257.

Kervinen, J., Tobin, G.J., Costa, J., Waugh, D.S., Wlodawer, A., Zdanov, A. (1999). Crystal structure of plant aspartic proteinase prophytepsin: inactivation and vacuolar targeting. *The EMBO journal*, 18(14): 3947–3955.

Kervinen, J., Tobin, G.J., Costa, J., Waugh, D.S., Wlodawer, A., Zdanov, A. (1999). Crystal structure of plant aspartic proteinase prophytepsin: inactivation and vacuolar targeting. *EMBO. J.*, 18: 3947–3955.

Kinoshita, T., Nishimura, M. and Hara-nishimura, I. (1995). Homologues of a vacuolar processing enzyme that are expressed in different organs in *Arabidopsis thaliana*. *Plant molecular biology*, 29: 81–89.

Kishimoto, Y., Hiraiwa, M. and O'Brien, J. S. (1992). Saposins: structure, function, distribution, and molecular genetics. *Journal of Lipid Research*, 33: 1255-1268.

Klein, A., Henseler, M., Klein, C., Suzuki, K., Harzer, K., Sandhoff, K. (1994). Sphingolipid activator protein D (sap-D) stimulates the lysosomal degradation of ceramide *in vivo*. *Biochem. Biophys. Res. Commun.*, 200:1440-1448.

Klimešová, J., Nobis, M.P., Herben, T. (2015). Senescence, ageing and death of the whole plant: morphological prerequisites and constraints of plant immortality. *New Phytol.* 206: 14–18.

Koelsch, G., Mares, M., Metcalf, P., and Fusek, M. (1994). Multiple functions of pro-peptides of aspartic proteinase zymogens. *Federation of European Biochemical Society*, 343: 6–10.

Konrad, S.S.A., Popp, C., Stratil, T.F., Jarsch, I.K., Thallmair, V., Folgmann, J., Marín, M.

and Ott, T. (2014). S-acylation anchors remorin proteins to the plasma membrane but does not primarily determine their localization in membrane microdomains. *New Phytol*, 203: 758-769.

Lagerholm, B.C., Weinreb, G.E., Jacobson, K., Thompson, N.L. (2005). Detecting microdomains in intact cell membranes. *Annu. Rev. Phys. Chem.*, 56:309–336

Laloi, M., McCarthy, J., Morandi, O., Gysler, C. and Bucheli, P. (2002). Molecular and biochemical characterization of two aspartic proteinases TcAP1 and TcAP2 from *Theobroma cacao* seeds. *Planta*, 215: 754–762.

Laloi, M., Perret, A.M., Chatre, L., Melser, S., Cantrel, C., Vaultier, M.N., Zachowski, A., Bathany, K., Schmitter, J.M., Vallet, M. et al. (2007). Insights into the role of specific lipids in the formation and delivery of lipid microdomains to the plasma membrane of plant cells. *Plant Physiol*, 143:461–472.

Leippe, M., Bruhn, H., Hecht, O., Grotzinger, J. (2005). Ancient weapons: the three-dimensional structure of amoebapore A. *Trends Parasitol* 21:5–7.

Lenarcic, T., Albert, I., Bohm, H., Hodnik, V., Pirc, K., Zavec, A.B., Podobnik, M., Pahovnik, D., Zagar, E., Pruitt R., et al. (2017). Eudicot plant-specific sphingolipids determine host selectivity of microbial NLP cytolysins. *Science*, 358: 1431-1434.

Leonova, T., Qi, X., Bencosme, A., Ponce, E., Sun, Y., and Grabowski, G.A. 1996. Proteolytic Processing Patterns of Prosaposin in Insect and Mammalian Cells. *J. Biol. Chem.*, 271: 17312-.

Li, X., Wang, X., Yang, Y., Li, R., He, Q., Fang, X., Lu, D.T., Maurel, C., Lin, J. (2011). Single-

molecule analysis of PIP₂;1 dynamics and partitioning reveals multiple modes of *Arabidopsis* plasma membrane aquaporin regulation. *The Plant Cell*, 23:3780–3797.

Liang, H., Yao, N., Song, J.T., Luo, S., Lu, H., Greenberg, J.T. (2003). Ceramides modulate programmed cell death in plants. *Genes and Development*, 17: 2636–2641.

Liepinsh, E., Andersson, M., Ruyschaert, J. et al. Saposin fold revealed by the NMR structure of NK-lysin. *Nat. Struct. Mol. Biol.* 4: 793–795.

Linke, T., Wilkening, G., Sadeghlar, F., Mozcall, H., Bernardo, K., Schuchman, E., Sandhoff, K. (2001). Interfacial regulation of acid ceramidase activity. Stimulation of ceramide degradation by lysosomal lipids and sphingolipid activator proteins. *J. Biol. Chem.*, 276: 5760–5768.

Lockshin, R.A., Zakeri, Z. (2004). Apoptosis, autophagy, and more. *The International Journal of Biochemistry and Cell Biology*, 36:2405-2419.

Lufrano, D., Faro, R., Castanheira, P., Parisi, G., Veríssimo, P., Vairo-Cavalli, S., Simões, I., Faro, C. (2012). Molecular cloning and characterization of procirsin, an active aspartic protease precursor from *Cirsium vulgare* (Asteraceae). *Phytochemistry*, 81: 7-18.

Lufrano, D., Faro, R., Castanheira, P., Parisi, G., Veríssimo, P., Vairo-Cavalli, S., Simões, I., Faro, C., Heimgartner, U., Pietrzak, M., Geertsen, R., Brodelius, P., da Silva Figueiredo, A.C., Pais, M.S.S. (1990). Purification and partial characterization of milk clotting proteases from flowers of *Cynara cardunculus*. *Phytochemistry*, 29(5): 1405-1410.

Lynch D.V., Dunn T.M. (2004). An introduction to plant sphingolipids and a review of

recent advances in understanding their metabolism and function. *New Phytologist*, 161: 677–702.

Lynch, D.V., Steponkus, P.L. (1987). Plasma membrane lipid alterations associated with cold acclimation of winter rye seedlings (*Secale cereale* L. cv Puma). *Plant Physiology*, 83: 761–767.

Markham, J.E., Jaworski, J.G. (2007). Rapid measurement of sphingolipids from *Arabidopsis thaliana* by reversed-phase high-performance liquid chromatography coupled to electrospray ionization tandem mass spectrometry. *Rapid Communications in Mass Spectrometry*, 21: 1304–1314.

Markham, J.E., Li, J., Cahoon, E.B., Jaworski, J.G. (2006). Separation and identification of major plant sphingolipid classes from leaves. *Journal of Bioorganic Chemistry*, 281: 22684–22694.

Markham, J.E., Molino, D., Gissot, L., Bellec, Y., Hematy, K., Marion, J., Belcram, K., Palauqui, J.C., Satiat-Jeunemaitre, B., Faure, J.D. (2011). Sphingolipids containing very-long-chain fatty acids define a secretory pathway for specific polar plasma membrane protein targeting in *Arabidopsis*. *The Plant Cell*, 23 (6):2362-2378.

Matsuda, J. (2008). Sphingolipid Activator Proteins. In: Taniguchi N., Suzuki A., Ito Y., Narimatsu H., Kawasaki T., Hase S. (eds) *Experimental Glycoscience*. Springer, Tokyo. pp. 125-129.

Matsuda, J., Vanier, M.T., Saito, Y., Tohyama, J., Suzuki, K., Suzuki, K. (2001). A mutation in the saposin A domain of the sphingolipid activator protein (prosaposin) gene results

in a late-onset, chronic form of globoid cell leukodystrophy in the mouse. *Human Molecular Genetics*, 10(11): 1191-1199.

Mazorra-Manzano, M.A., Yada, R.Y. (2008). Expression and characterization of the recombinant aspartic proteinase A1 from *Arabidopsis thaliana*. *Phytochemistry*, 69(13): 2439-2448

Mendieta, J.R., Pagano, M.R., Munoz, F.F., Daleo, G.R. & Guevara, M.G. (2006). Antimicrobial activity of potato aspartic proteases (StAPs) involves membrane permeabilization. *Microbiology*, 152: 2039-2047.

Michaelson, L.V., Zauner, S., Markham, J.E., Haslam, R.P., Desikan, R., Mugford, S., Albrecht, S., Warnecke, D., Sperling, P., Heinz, E., et al. (2009). Functional characterization of a higher plant sphingolipid Delta4-desaturase: defining the role of sphingosine and sphingosine-1-phosphate in *Arabidopsis*. *Plant Physiology*, 149: 487-498.

Minami, A., Fujiwara, M., Furuto, A., Fukao, Y., Yamashita, T., Kamo, M., Kawamura, Y., Uemura, M. (2009). Alterations in detergent-resistant plasma membrane microdomains in *Arabidopsis thaliana* during cold acclimation. *Plant and Cell Physiology*, 50: 341-359.

Misas-Villamil, J.C., Toenges, G., Kolodziejek, I., Sadaghiani, A.M., Kaschani, F., Colby, T., Bogyo, M. and Van der Hoorn, R.A.L. (2013). Activity profiling of vacuolar processing enzymes reveals a role for VPE during oomycete infection. *The Plant journal*, 73(4): 689–700.

Mochizuki, S., Harada, A., Inada, S., Sugimoto-Shirasu, K., Stacey, N., Wada, T., Ishiguro, S., Okada, K., Sakai, T. (2005). The *Arabidopsis* WAVY GROWTH 2 Protein Modulates Root Bending in Response to Environmental Stimuli. *The Plant Cell* Feb, 17 (2): 537-547.

Molino, D., Van der Giessen, E., Gissot, L., Hematy, K., Marion, J., Barthelemy, J., Bellec Y., Vernhettes, S., Satiat-Jeunemaitre, B., Galli, T., Taresté, D., Fauré J.D. (2014). Inhibition of very long acyl chain sphingolipid synthesis modifies membrane dynamics during plant cytokinesis. *Biochim. Biophys. Acta.*, 1842 (10):1422-1430.

Mongrand, S., Morel, J., Laroche, J., Claverol, S., Carde, J.P., Hartmann, M.A., Bonneau, M., Simon-Plas, F., Lessire, R., Bessoule, J.J. (2004). Lipid rafts in higher plant cells: purification and characterization of Triton X-100-insoluble microdomains from tobacco plasma membrane. *J. Biol. Chem.*, 279:36277–36286.

Morimoto, S., Kishimoto, Y., Tomich, J., Weiler, S., Ohashi, T., Barranger, J. A., Kretz, K. A., O'Brien, J. S. (1990). Interaction of saposins, acidic lipids, and glucosylceramidase. *J. Biol. Chem.*, 265:1933-1937

Morimoto, S., Martin, B. M., Yamamoto, Y., Kretz, K. A., O'Brien, J. S., Kishimoto, Y. (1989). Saposin A: second cerebroside activator protein. *Proc. Natl. Acad. Sci. U.S.A.*, 86: 3389-3393.

Muñoz, F., Mendieta, J., Pagano, M., Paggi, R., Daleo, G., Guevara, M. (2010). The swaposin-like domain of potato aspartic protease (StAsp-PSI) exerts antimicrobial activity on plant and human pathogens. *Peptides*. 31: 777-85.

- Muñoz, F., Palomares-Jerez, M.F., Daleo, G., Villalaín, J., Guevara, M.G. (2011). Cholesterol and membrane phospholipid compositions modulate the leakage capacity of the swaposin domain from a potato aspartic protease (StAsp-PSI). *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids*, 1811(12): 1038-1044.
- Muñoz, F. F., Mendieta, J. R., Pagano, M. R., Paggi, R. A., Daleo, G. R. & Guevara, M. G. (2010). The swaposin-like domain of potato aspartic protease (StAsp-PSI) exerts antimicrobial activity on plant and human pathogens. *Peptides*, 31(5): 777-785.
- Muñoz, F., Palomares-Jerez, M. F., Daleo, G., Villalaín, J. & Guevara, M. G. (2014). Possible mechanism of structural transformations induced by StAsp-PSI in lipid membranes. *Biochimica et Biophysica Acta - Biomembranes*, 1838(1, Part B): 339-347.
- Müntz, K. and Muntz, K. (1998). Deposition of storage proteins. *Plant molecular biology*, 38(1-2): 77–99.
- Murakami, S., Kondo, Y., Nakano, T., (2000). Protease activity of CND41, a chloroplast nucleoid DNA-binding protein, isolated from cultured tobacco cells. *FEBS Letters* 468: 15-18.
- Mutlu, A., Chen, X., Reddy, S.M. and Gal, S. (1999). The aspartic proteinase is expressed in *Arabidopsis thaliana* seeds and localized in the protein bodies. *Seed science research*, 9(1): 75–84.
- Mutlu, A. and Gal, S. (1999). Plant aspartic proteinases: enzymes on the way to a function. *Physiologia Plantarum*, 5: 569–576.
- Nakano, T., Murakami, S., Shoji, T., Yoshida, S., Ymada, Y., Sata, F., (1997). A novel

protein with DNA binding activity from tobacco chloroplast nucleoids. *The Plant Cell* 9: 1673-1682.

Nakaune, S., Yamada, K., Kondo, M., Kato, T., Tabata, S. and Nishimura, M. (2005). A vacuolar processing enzyme, δ VPE, is involved in seed coat formation at the early stage of seed development. *The Plant cell*, 17(3): 876–887.

O'Brien, J.S., Kishimoto, Y. (1991). Saposin protein: structure, function, and role in human lysosomal storage disorder. *Official Journal of the Federation of American Societies for Experimental Biology*, 5: 301-308.

Olmeda, B., García-Álvarez, B. & Pérez-Gil, J. (2013). Structure–function correlations of pulmonary surfactant protein SP-B and the saposin-like family of proteins. *Eur. Biophys. J.* 42: 209–222.

Olvera-Carrillo, Y., Bel, M.V., Hautegeem, T.V., Fendrych, M., Huysmans, M., Simaskoya M., Durme, M.V., Buscaill, P., Rivas, S., Coll, N.S., Coppens, F., Maere, S., Nowack, M.K. (2015). A conserved core of programmed cell death indicator genes discriminates developmentally and environmentally induced programmed cell death in plants. *Plant Physiology*, 169 (4): 2684-2699.

Olvera-Carrillo, Y., Van Bel, M., Van Hautegeem, T., Fendrych, M., Huysmans, M., et al. (2015). A conserved core of programmed cell death indicator genes discriminates developmentally and environmentally induced programmed cell death in plants. *Plant Physiol.* 169: 2684-2699.

Olvera-Carrillo, Y., Bel, MV., Hautegeem, TV., Fendrych, M., Huysmans, M., Simaskoya,

M., Durme, MV., Buscaill, P., Rivas, S., Coll, NS., Coppens, F., Maere, S., Nowack, MK. (2015). A conserved core of programmed cell death indicator genes discriminates developmentally and environmentally induced programmed cell death in plants. *Plant Physiology*, 169 (4): 2684-2699.

Otegui, M.S., Herder R., Schulze, J., Jung, R. and Staehelin, L.A. (2006). The proteolytic processing of seed storage proteins in *Arabidopsis* embryo cells starts in the multivesicular bodies. *The Plant Cell*, 18(10): 2567–2581.

Otegui, MS., Herder, R., Schulze, J., Jung, R., Staehelin, LA. (2006). The proteolytic processing of seed storage proteins in *Arabidopsis* embryo cells starts in the multivesicular bodies. *The Plant Cell*, 18(10):2567-2581.

Pagano, M. R., Mendieta, J. R., Muñoz, F. F., Daleo, G. R. & Guevara, M. G. (2007). Roles of glycosylation on the antifungal activity and apoplast accumulation of StAPs (*Solanum tuberosum* aspartic proteases). *International Journal of Biological Macromolecules*, 41(5): 512-520.

Paris, N., Stanley, C.M., Jones, R.L. and Rogers, .JC. (1996). Plant cells contain two functionally distinct vacuolar compartments. *Cell*, 85(4): 563–572.

Pata M.O., Hannun Y.A., Ng C.K. (2010). Plant sphingolipids: decoding the enigma of the Sphinx. *New Phytologist*, 185: 611–630.

Pena, S.V., and Krensky, A. M. (1997). Granulysin, a new human cytolytic granuleassociated protein with possible involvement in cell-mediated cytotoxicity. *Semin. Immunol.* 9: 117–125.

- Pena, S.V. and Krensky, A.M. (1997). Granulysin, a new human cytolytic granule-associated protein with possible involvement in cell-mediated cytotoxicity. *Semin. Immunol.*, 9: 117–125
- Pereira, C., Pereira S., Satiat-Jeunemaitre, B. and Pissarra, J. (2013). Cardosin A contains two vacuolar sorting signals using different vacuolar routes in tobacco epidermal cells. *The Plant Journal*, 76(1): 87-100.
- Pereira, C.S., de Costa, D.S., Pereira, S., Nogueira, F.M. Albuquerque, P.M., Teixeira, J., Faro, C., Pissarra, J. (2008). Cardosins in postembryonic development of cardoon: towards an elucidation of the biological function of plant aspartic proteinases. *Protoplasma* 232: 203–213.
- Petrov, V., Hille J., Mueller-Roeber, B., Gechev, T.S. (2015). ROS-mediated abiotic stress-induced programmed cell death in plants. *Front. Plant Sci.* 6: 69.
- Piatigorsky, J., Norman, B., Dishaw, L. J., Kos, L., Horwitz, J., Steinbach, P. J. and Kozmik, Z. (2001). J3-crystallin of the jellyfish lens: similarity to saposins. *Proc. Natl. Acad. Sci.*, 98: 12362–12367.
- Popovic, K., Holyoake, J., Pomès, R., Privé, G.G. (2012). Structure of saposin a lipoprotein discs. *Proc. Natl. Acad. Sci. USA*, 109: 2908–2912.
- Popovic, K., Privé, G.G. (2008). Structures of the human ceramide activator protein saposin D. *Acta Crystallogr. D*, 64: 589–594.
- Popovic, K., Privé, G.G. (2008). Structures of the human ceramide activator protein saposin D. *Acta Crystallogr. D*, 64: 589–594.

Raimbault, A., Zuily-Fodil, Y., Soler, A., Cruz de Carvalho, M.H. (2013). A novel aspartic acid protease gene from pineapple fruit (*Ananas comosus*): Cloning, characterization and relation to postharvest chilling stress resistance. *Journal of Plant Physiology*, 170(17): 1536-1540.

Reape, T.J., Molony, E.M., McCabe, P.F. (2008). Programmed cell death in plants: distinguishing between different modes. *Journal of Experimental Botany*, 59(3):435–444.

Richardt, S., Lang, D., Reski, R., Frank, W., & Rensing, S.A. (2007). PlantAPDB, a phylogeny-based resource of plant transcription-associated proteins. *Plant physiology*, 143(4): 1452–1466.

Rossmann, M., Schultz-Heienbrok, R., Behlke, J., Rimmel, N., Alings, C., Sandhoff, K., Saenger, W., Maier, T. (2008). Crystal Structures of Human Saposins C and D: Implications for Lipid Recognition and Membrane Interactions. *Structure*, 16(5): 809-817.

Rossmann, M., Schultz-Heienbrok, R., Behlke, J., Rimmel, N., Allings, C., Sandhoff, K., Saenger, W., Maier, T. (2008). Crystal structures of human saposins C and D: Implications for lipid recognition and membrane interactions. *Structure*, 16: 809–817.

Sansen, S., De Ranter, C.J., Gebruers, K., Brijs, K., Courtin, C.M., Delcour, J.A., Rabijns, A. (2004). Structural Basis for Inhibition of *Aspergillus niger* Xylanase by *Triticum aestivum* Xylanase Inhibitor-I J. *Biol. Chem.*, 279: 36022-.

Sarkkinen, P., Kalkkinen, N., Tilgmann, C. et al. (1992). Aspartic proteinase from barley

grains is related to mammalian lysosomal cathepsin D. *Planta* 186: 317–323.

Sarkkinen, P., Kalkkinen, N., Tilgmann, C. et al. (1992). Aspartic proteinase from barley grains is related to mammalian lysosomal cathepsin D. *Planta*, 186: 317–323.

Sarmiento, A.C., Lopes, H., Oliveira, C.S., Vitorino R., Samyn B., Sergeant K., Debyser G., Beeumen, J.V. Domingues, P., Amado, F., Pires, E., Domingues, M.R.M., Barros, M.T. (2009). Multiplicity of aspartic proteinases from *Cynara cardunculus* L.. *Planta* 230: 429–439.

Schaaf, A., Reski, R., Decker, E.L. (2004). A novel aspartic proteinase is targeted to the secretory pathway and to the vacuole in the moss *Physcomitrella patens*. *European Journal of Cell Biology*, 83(4):145-152

Schuetz, C. G., Pierstorff, B., Huettler, S. and Sandhoff, K. (2001). Sphingolipid activator proteins: proteins with complex functions in lipid degradation and skin biogenesis. *Glycobiology*, 11: 81R–90R.

Shirakawa, M., Ueda, H., Shimada, T., Koumoto, Y., Shimada, T.L., Kondo, M., Takahashi, T., Okuyama, Y., Nishimura, M. and Hara-Nishimura, I. (2010). *Arabidopsis* Qa-SNARE SYP2 proteins localized to different subcellular regions function redundantly in vacuolar protein sorting and plant development. *The Plant Journal*, 64: 924-935.

Simoës, I., Faro, C. (2004). Structure and function of plant aspartic proteinases. *European Journal of Biochemistry*, 271: 2067-2075

Simoës, I. & Faro, C. (2004). Structure and function of plant aspartic proteinases. *The FEBS Journal*, 271(11): 2067-2075.

Spassieva, S.D., Markham, J.E., Hille, J. (2002). The plant disease resistance gene Asc-1 prevents disruption of sphingolipid metabolism during AAL-toxin-induced programmed cell death. *Plant Journal*, 32: 561-572.

Spiegel, R., Bach, G., Sury, V., Mengistu, G., Meidan, B., Shalev, S., Shneor, Y., Mandel, H., Zeigler, M. (2005). A mutation in the saposin A coding region of the prosaposin gene in an infant presenting as Krabbe disease: first report of saposin A deficiency in humans. *Molecular Genetics and Metabolism*, 84(2):160-166.

Staab, J. F., Ginkel, D. L., Rosenberg, G. B. and Munford, R. S. (1994). A saposin-like domain influences the intracellular localization, stability, and catalytic activity of human acyloxyacyl hydrolase. *J. Biol. Chem.*, 269: 23736–23742

Stael, S., Breusegem, F.V., Gevaert, K., Nowack, M.K. (2019). Plant proteases and programmed cell death. *Journal of Experimental Botany*, 70:1991–1995.

Swanson, S., Bethke, P. and Jones, R. (1998). Barley aleurone cells contain two types of vacuoles: characterization of lytic organelles by use of fluorescent probes. *The Plant cell*, 10(5): 685–698.

Takahashi, D., Imai, H., Kawamura, Y., Uemura, M. (2016). Lipid profiles of detergent resistant fractions of the plasma membrane in oat and rye in association with cold acclimation and freezing tolerance. *Cryobiology*, 72: 123–134.

Takahashi, D., Imai, H., Kawamura, Y., Uemura, M. (2016). Lipid profiles of detergent resistant fractions of the plasma membrane in oat and rye in association with cold acclimation and freezing tolerance. *Cryobiology*, 72: 123–134.

- Tamargo, R.J., Velayati, A., Goldin, E., Sidransky, E. (2012). The role of saposin C in Gaucher disease. *Molecular Genetics and Metabolism*, 106(3):257-263.
- Tayama, M., Soeda, S., Kishimoto, Y., Martin, B. M., Callahan, J. W., Hiraiwa, M., O'Brien, J. S. (1993). Effect of saposins on acid sphingomyelinase. *Biochem. J.*, 290: 401-404
- Terauchi, K., Asakura, T., Ueda, H., Tamura, T., Tamura, K., Matsumoto, I., Misaka, T., Hara-Nishimura, I., Abe, K. (2006). Plant-specific insertions in the soybean aspartic proteinases, soyAP1 and soyAP2, perform different functions of vacuolar targeting. *Journal of Plant Physiology*, 163(8): 856-862.
- Titapiwatanakun, B., Blakeslee, J.J., Bandyopadhyay, A., Yang, H., Mravec, J., Sauer, M., Cheng, Y., Adamec, J., Nagashima, A., Geisler, M., Sakai, T., Friml, J., Peer, W.A. and Murphy, A.S. (2009). ABCB19/PGP19 stabilizes PIN1 in membrane microdomains in *Arabidopsis*. *The Plant Journal*, 57: 27-44.
- Tormakangas, K., Hadlington, J.L., Pimpl, P., Hillmer, S., Brandizzi, F., Teeri, T.H. & Denecke, J. (2001). A vacuolar sorting domain may also influence the way in which proteins leave the endoplasmic reticulum. *The Plant Cell*, 13(9): 2021-2032.
- Tse, Y.C., Wang, J., Miao, Y. and Jiang, L. (2007). Biogenesis of the compound seed protein storage vacuole. *In Seeds: biology, development and ecology*. pp. 112–119.
- Vaccaro, A.M., Ciaffoni, F., Tatti, M., Salvioli, R., Barca, A., Tognozzi, D., Scerch, C. (1995). pH-dependent conformational properties of saposins and their interactions with phospholipid membranes. *J. Biol. Chem.*, 27: 30576–30580.
- Vieira, M., Pissarr, J., Veríssimo, P., Castanheira, P., Costa, Y., Pires, E., Faro, C. (2001).

Molecular cloning and characterization of cDNA encoding cardosin B, an aspartic proteinase accumulating extracellularly in the transmitting tissue of *Cynara cardunculus* L.. *Plant Mol Biol* 45: 529–539.

Vieira, M., Pissarra, J., Veríssimo, P., Castanheira, P., Costa, Y., Pires, E. & Faro, C. (2001).

Molecular cloning and characterization of cDNA encoding Cardosin B, an aspartic proteinase accumulating extracellularly in the transmitting tissue of *Cynara cardunculus* L. *Plant Molecular Biology*, 45(5): 529-539.

Vieira, V., Peixoto, B., Costa, M., Pereira, S., Pissarra, J., Pereira, S. (2019). N-linked glycosylation modulates Golgi-independent vacuolar sorting mediated by the plant specific insert. *Plants*, 8(9):312.

Vieira, V., Peixoto, B., Costa, M., Pereira, S., Pissarra, J., Pereira, C. (2019). N-Linked glycosylation modulates Golgi-independent vacuolar sorting mediated by the plant specific insert. *Plants*, 8: 312.

Vogel, A., Schwarzmann, G., Sandhoff, K. (1991). Glycosphingolipid specificity of the human sulfatide activator protein. *Eur. J. Biochem.*, 200: 591-597

Warnecke, D., Heinz, E. (2003). Recently discovered functions of glucosylceramides in plants and fungi. *Cellular and Molecular Life Sciences*, 60: 919–941.

Wenger, D. A., Sattler, M., Roth, S. (1982). A protein activator of galactosylceramide β -galactosidase. *Biochim. Biophys. Acta*, 712: 639-649

Wilkins, K.A., Poulter, N.S., Franklin-Tong, V.E. (2014). Taking one for the team: self-recognition and cell suicide in pollen. *J. Exp. Bot.* 65: 1331–1342.

- Winau, F., Schwierzeck, V., Hurwitz, R., Rimmel, N., Sieling, P. A., Modlin, R. L., Porcelli S. A., Brinkmann, V., Sugita, M., Sandhoff, K. et al. (2004). Saposin C is required for lipid presentation by human CD1b. *Nat. Immunol.* 5: 169–174.
- Winter, D., Vinegar, B., Nahal, H., Ammar, R., Wilson, G.V. and Provart, N.J. (2007). An “Electronic Fluorescent Pictograph” browser for exploring and analyzing large-scale biological data sets. *PLoS ONE*, 2(8), e718.
- Worrall, D., Ng, C.K.Y., Hetherington, A.M. (2003). Sphingolipids, new players in plant signaling. *Trends in Plant Science*, 8: 317–320.
- Wrobel, R., Jones, B.L. (1992). Appearance of Endoproteolytic Enzymes during the Germination of Barley. *Plant Physiology*, 100 (3): 1508-1516.
- Wrobel, R., Jones, B.L. (1992). Appearance of endoproteolytic enzymes during the germination of barley. *Plant Physiology*, 100 (3) :1508-1516.
- Wu, L., Chen, H., Curtis, C., Fu, Z.Q. (2014). Go in for the kill: how plants deploy effector-triggered immunity to combat pathogens. *Virulence*, 5: 710–721.
- Xiang, L., Etxeberria, E. and Van den Ende, W. (2013). Vacuolar protein sorting mechanisms in plants. *FEBS journal*, 280(4), 979–993.
- Xuan, W., Band, L.R., Kumpf, R.P., Van Damme, D., Parizot, B., De Rop, G., Opdenacker, D., Möller, B.K., Skorzinski, N., Njo, M.F., De Rybel, B., Audenaert, D., Nowack, M.K., Vanneste, S., Beeckman, T. (2016). Cycles of programmed cell death establish the developmental clock in plant roots. *Science*, 351: 384-387.
- Yang, H., Richter, G.L., Wang, X., Młodzieńska, E., Carraro, N., Ma, G., Jenness, M., Chao,

- D.-Y., Peer, W.A. and Murphy, A.S. (2013). Sterols and sphingolipids differentially function in trafficking of the *Arabidopsis* ABCB19 auxin transporter. *Plant. J.*, 74: 37-47.
- Yao X., Xiong W., Ye T., Wu Y. (2012). Overexpression of the aspartic protease ASPG1 gene confers drought avoidance in *Arabidopsis*. *Journal of Experimental Botany*, 63(7):2579–2593.
- Ye, H., Ren, F., Guo, H., Guo, L., Bai, J., Wang, Y. (2020). Identification of key genes and transcription factors in ageing *Arabidopsis* papilla cells by transcriptome analysis. *Plant Physiology and Biochemistry*, 147:1-9.
- Ye, H., Ren, F., Guo, H., Guo, L., Bai, J., Wang, Y. (2020). Identification of key genes and transcription factors in ageing *Arabidopsis* papilla cells by transcriptome analysis. *Plant Physiology and Biochemistry*, 147:1-9.
- Yu, M., Cui, Y., Zhang, X. et al. (2020). Organization and dynamics of functional plant membrane microdomains. *Cell. Mol. Life Sci.*, 77: 275–287.
- Yuan, W., Qi, X., Tsang, P., Kang, S.J., Ilarionov, P.A., Besra, G.S., Gumperz, J., Cresswell, P. (2007). Saposin B is the dominant saposin that facilitates lipid binding to human CD1d molecules. *Proc. Natl. Acad. Sci. USA*, 104: 5551–5556.
- Zhou, D., Cantu 3rd, C., Sagiv, Y., Schrantz, N., Kulkarni, A. B., Qi, X., Mahuran, D. J., Morales, C. R., Grabowski, G. A., Benlagha, K. et al. (2004). Editing of CD1d-bound lipid antigens by endosomal lipid transfer proteins. *Science*, 303: 523–527.
- Zouhar, J., Munoz, A. and Rojo, E. (2010). Functional specialization within the vacuolar sorting receptor family: VSR1, VSR3 and VSR4 sort vacuolar storage cargo in seeds and

vegetative tissues. *Plant J.* 64: 577–588.